National Library of Canada
Canadian Theses Service
Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Service des thèses canadiennes

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

7

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-55364-2

Canada

THE UNIVERSITY OF ALBERTA

SANDEL ON RAWLS AND THE LIMITS OF LIBERALISM

BY

JILL HUNTER

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF ARTS

DEPARTMENT OF PHILOSOPHY

EDMONTON, ALBERTA

FALL, 1989

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR Jill Hunter

TITLE OF THESIS Sandel on Rawls and the Limits of Liberalism

DEGREE FOR WHICH THESIS WAS PRESENTED MASTER OF ARTS

YEAR THIS DEREE GRANTED Fall, 1989

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED). Jill Hunter

PERMANENT ADDRESS:

9028-101A Avenue

Edmonton

Alberta

Date October 5 19 89

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "Sandel on Rawls and The Limits of Liberalism" submitted by Jill Hunter in partial fulfillment of the requirements for the degree of Master of Arts.

.....
Supervisor

.....

.....

.....

Date: August 10 ,1989.

ABSTRACT

The dispute between liberals and communitarians is at the fore in contemporary political philosophy. In Liberalism and the Limits of Justice, Michael Sandel's unique criticism of Rawlsian liberalism has directed the discussion towards metaphysics and ontology as opposed to merely advocating a particular view of society and the preferred role of government. I argue that despite Sandel's original approach his critique of Rawls' A Theory of Justice is unsuccessful. These new issues have prompted Charles Taylor, a communitarian, to comment on what he views as an obstacle to the progress of such a discussion. In his view, the different issues being addressed must be distinguished. Specifically, he charges liberals with failing to distinguish between the view of society that communitarians advocate as their own and communitarian critiques of liberal metaphysics and ontology. However, in this particular case Sandel himself grounds his critique of Rawlsian liberalism on the conflation of metaphysics and ethics. Sandel wishes to attribute two incompatible metaphysical conceptions of the self to Rawls. Yet Sandel's arguments are based on a failure to distinguish between issues concerning metaphysics, epistemology, ethics and politics. Even within metaphysics his critique of Rawls depends on conflating questions involving personal identity with

questions involving individuation. Once these distinctions are made the fallacious character of Sandel's arguments are revealed and his critique of Rawlsian metaphysics disappears.

PREFACE

Following the publication of Michael Sandel's Liberalism and the Limits of Justice, the debate between communitarians and liberals is shifting from issues in political philosophy to issues concerning metaphysics and the philosophy of mind. Communitarians defend the political thesis that government and society should promote a monistic conception of the good life. Among variations in these views is a common neo-Aristotelian notion of human excellence or human flourishing. Moreover, they believe that human excellence can only be realized within a community. Some communitarians believe that cohesive monistic views of the good can be found in existing societies. While others believe that such views will be adopted by reasonable people once the shared basis for their attachments to their communities is revealed. On the other side of the debate liberals point to pluralism and a commitment to rights of the individual as factors limiting the role of government. In cases where justice is not an issue, they argue government must remain neutral with respect to the good because society is composed of individuals some of whom hold incommensurable conceptions of the good. Liberals advocate a limited role for the government because conflicting conceptions of the good often cannot be reconciled without the oppressive use of state power and the violation of individual rights.

Sandel's novel approach to the communitarian/liberal debate attacks Rawls' liberalism by going beyond a political critique and challenging the metaphysics behind the theory. I shall argue that this approach to the critique of A Theory of Justice is unsuccessful and must be dismantled. This is what I propose to accomplish here.

Sandel argues that Rawls' theory is supported by two incompatible metaphysical conceptions of the self. He maintains that if Rawls wishes to retain the deontological character of his theory and his view of distributive justice, then A Theory of Justice presupposes the Kantian self and the constitutive conception of the self. Sandel attributes his interpretation of the Kantian ego to Rawls on the following grounds: If the principles of justice are derived independently of conceptions of the good life (i.e., if his theory is deontological) then Rawls' theory presupposes a self in which all empirical properties are contingent to personal identity. Second, he argues that if Rawls upholds his view of distributive justice -- represented by the claim, 'social and economic inequalities must be to the greatest benefit of the least advantaged members of society', -- then the constitutive conception of the self is entailed. The constitutive conception of the self involves viewing the community as an artificial entity that possesses properties normally viewed as attributes of an empirically individuated self (i.e., a person).

Initially Sandel's criticisms seem unavoidable. Yet his attempt to ground A Theory of Justice in metaphysics and the philosophy of mind is based on a misreading of Rawls. Rawls may require some metaphysical foundation for his theory but Sandel neglects to consider more plausible alternatives that could serve this purpose. Sandel refuses to acknowledge the extent to which Rawls owes methodological debts to thinkers like Quine. For Quine, science determines what counts as good explanation or justification in the scientific community, and similarly for Rawls the political community determines what counts as relevant and good reasons for accepting a conception of justice. Rawls looks to Quine insofar as the method for determining the principles of justice must be publicly accepted by potential citizens. In short, Rawls is a pragmatist. Moreover, he is not committed to defending Quine's views of justification and truth as philosophical truths because A Theory of Justice does not require such grounding. Rawls sidesteps controversial philosophical issues generally, and controversial metaphysical issues specifically by deriving the principles of justice from a basis of public agreement. Issues that cannot be publicly resolved upon reflection cannot be grounds for accepting a conception of justice. By exposing Sandel's errors I provide a clearer and more accurate account of A Theory of Justice.

ACKNOWLEDGEMENTS

I wish to thank my thesis supervisor Wes Cooper for his constant encouragement, patients, and insightful criticisms. I am also indebted to Steven DeHaven for suggesting ways of expressing my claims more rigorously, as a result of which this project became clearer than it might have otherwise been. Lastly, I am grateful to my husband, Bruce Hunter who spent many hours with me discussing the issues between communitarians and liberals. Without this daily feedback many problems would have remained unresolved.

Table of Contents

Chapter Page

I. Introduction 1

1. The Primacy of Justice 9

A. Justice as a First Virtue 11
B. Justice and Procedure 25

II. 51

1. Logical, Conceptual, and Epistemological Priority.....51

2. The Kantian Self 54

3. The Priority of the Rawlsian Subject 66

4. Individuation and Personal Identity.....67

5. The Epistemology and the Metaphysics of the Self.....72

III. 78

1. The Kantian Ego/Fallaciously Derived 80

2. The Constitutive Conception of the Self/ Metaphyscial or Moral 89

3. The Constitutive Conception/Fallaciously Derived.....99

4. Rawls' Reply to the Dilemma 104

5. Objectivism-The Difference between Rawls and Sandel.....109

Bibliography 121

CHAPTER I

Introduction

In Liberalism and the Limits of Justice, Michael Sandel develops a communitarian critique of the dominant strain of liberal thought according to which, justice is the first virtue of social institutions. Crucial to the 'deontological' liberalism, which Sandel takes to be represented by the work of John Rawls, is the idea that principles of justice are derived independently of conceptions of the good. The principles of justice are designed to regulate and resolve conflicts over social goods. Sandel's alternative conception of the ideal society is one in which the social virtues of benevolence and fraternity are at least as central as justice and, by reducing conflict, reduce the role or need for principles of justice. Sandel's attack on deontological liberalism goes much deeper, however. Beyond merely advocating his particular view of the good he presents two dilemmas for Rawls that question the metaphysics Sandel believes grounds A Theory of Justice.[1] First, he argues that if Rawlsian theory is truly one that derives the principles of justice independently of conceptions of the good then

---

[1]At a recent lecture at the University of Alberta, Charles Taylor reveals one of the main problems with the communitarian/liberal debate. Taylor claims that critics of communitarianism fail to separate two main issues. The first issue is the position that communitarians advocate as their own and the second issue concerns a critique of metaphysics and ontology that communitarians see as the foundation for deontological liberalism.

the hypothetical constructs (the people) in the original position have no ties to the empirical world. In short, he believes that the constraints that are imposed by the veil of ignorance sever any empirical foundation of the parties. Hence, if Rawls wishes to maintain the deontological character of his theory Sandel believes that he must think of the people in the original position as Kantian disembodied subjects.1 On the other hand, if this abstract metaphysically dubious conception of the self is not the foundation of Rawls' theory Sandel contends it fails as deontology and must presuppose conceptions of the good.

Second, Sandel argues that Rawls' view of distributive justice (treating assets and talents as common assets) presupposes a constitutive conception of the self according to which individuals within a community identify and share the values and goals of that community. However, this is incompatible with Rawls' individualistic project. Hence, we are introduced to a second dilemma. Either Rawls must reject the difference principle or he must adopt the constitutive conception of the self and abandon his vision of liberalism. An important aspect of Sandel's general argument is his methodology. Sandel arrives at these two conceptions of the self via a transcendental argument of sorts. He takes Rawls' political theory as given and argues back to

---

1Michael Sandel, Liberalism and the Limits of Justice, (New York: Cambridge University Press), p. 14.

the claim that A Theory of Justice must presuppose two distinct views of the self.1 Sandel introduces his criticisms as two dilemmas both of which appear to be damaging to Rawls. Hence, even if I can show that the metaphysical conceptions of the self do not follow from Rawls view I must also show that Rawls is not forced to accept the alternate horns of both dilemmas. My interpretation of Rawls opens a third alternative to the

---

1Sandel justifies these views of the self via a transcendental argument. He begins with Rawls' theory as his premise and tries to show that this entails a particular conception of the self. Sandel concedes that the conceptions of the self he accredits to Rawls are not explicit in A Theory of Justice, but he is convinced that his interpretation of Rawls' theory is confirmed through the "...hints and traces of evidence scattered throughout the text and, more importantly, in the sense it enables us to make of Rawls' theory as a whole,..."(Sandel, p. 47) Sandel's strategy and its justification, however, are firmly rooted in the text. The original position and reflective equilibrium provide the working material from Rawls' account to launch his general project.(Sandel, pp. 47-48)

Sandel's strategy is justified by appealing to the dual feature of reflective equilibrium. Sandel claims that if reflective equilibrium involves our considered convictions about justice on the one hand, and our judgments of 'reasonableness and plausibility' on the other hand, then the original position must also yield a product that is two-fold. Rawls' theory of justice is only one feature of the product that results from the original position. Sandel is convinced that a conception of the subject is foundational to a theory of justice because before we can have a theory of justice there must be subjects that affirm the theory. More importantly, the theory of the subject must be a theory of a certain kind; it must be a theory of a "moral subject". In effect, Sandel is claiming that a theory of justice presupposes a normative conception of the subject.

Given Rawls' own account of how reflective equilibrium operates, Sandel's strategy is prima facie sound. The principles of justice are given (which includes ethical claims about the moral arbitrariness of our assets) as premises in Sandel's arguments whereby the two conceptions of the self are seen to be entailed as conclusions.(Sandel, p. 49)

avenues suggested by Sandel. I argue that these views of the self do not follow from Rawls' theory. Moreover the dilemma Sandel thinks follows - no deontology or no difference principle - rests on misguided interpretation of A Theory of Justice. However, before this can be accomplished a brief introduction to Rawls' theory is necessary.

A Theory of Justice as now understood by John Rawls, is an attempt to find a rational solution to the practical problem of social justice in a modern liberal democracy.1

1In later articles, Rawls claims that A Theory of Justice is designed for a modern constitutional democracy. However, opponents of the theory may argue that on this point, Rawls has merely changed his view about what his book was designed to accomplish. However, we can see whether A Theory of Justice is compatible with such an interpretation which will lend support for Rawls' later claims. First, Rawls distinguishes between the special and the general conceptions of justice. In the former the two principles of justice are lexically ordered so that liberty cannot be traded for economic advantage whereas this is not so for the latter conception. Second, in the part two of the book Rawls applies the principles of justice to the institutions of a constitutional democracy. On page 195 of his book he claims that his intention is to apply the principles to a basic structure that satisfies those principles and those institutions are those of a democracy. Further, he states that "I do not argue that these arrangements are the only ones that are just. Rather my intention is to show that the principles of justice,..., define a workable political conception, and are a reasonable approximation to and extension of our considered judgments." The sense in which Rawls is a pragmatist is important here. What we can say about the theory is that it is clear from the book that the principles of justice do apply to a modern constitutional democracy. But whether the principles apply to other societies in which its members have different intuitions about justice is an open question that the book was not designed to resolve. Moreover, even if the special conception of justice cannot be applied to other societies it is plausible to suppose that some variation of the general conception of justice would apply but this again is an open question that cannot be resolved until the reasoning process is simulated including the different variables. As

The principles of justice are the result of a reasoning process devised by Rawls. The reasoning procedure is designed in such a way as to ensure that the principles that are chosen are the most reasonable and fair given that individuals do not agree on conceptions of the good. The content of the two principles define the terms of distributing liberties, duties and social goods. The principles are used for evaluating and the reforming the basic institutions of society. The theory is representative of deontological liberalism insofar as the principles of justice (the principles of right) are arrived at independently of a comprehensive view of the good life.2 In order to achieve this, the reasoning process is represented as a hypothetical contract in which the parties come to an agreement on the principles

---

a pragmatist Rawls relies on a coherence view of justification. By so doing, Rawls does not need a priori arguments which depend on viewing principles as necessary truths. Hence, in order to get the discussion started Rawls appeals to intuitions about justice and conditions for rationality that can be publicly accepted. Therefore, if these are accepted by other societies there is no reason to suppose that they would not also choose the principles of justice, all other things being equal. Even though I believe that the most favorable interpretation of A Theory of Justice is that it is intended to apply to a modern constitutional democracy however, I am willing to accept that perhaps Rawls' position has changed and that he should have revealed this limitation of his theory.

2A deontological theory is distinguished from a teleological theory on the grounds that in the former, the principles of justice and the right are derived independently of conceptions of the good. The latter, however, reduce the right to whatever they view to be the good.

of justice while adhering to certain constraints.3 Rawls believes that the reasoning process dramatized by the original position reflects the idea of justice as fairness.

3Jean Hampton argues that the original position not only need not but more importantly cannot be viewed as a contract. I agree with Hampton insofar as I think it's clear that the original position need not be viewed as a contract but could instead be viewed strictly as a reasoning process which could be simulated by one individual. However, I have reservations about her arguments that Rawls could not describe the original position as a social contract theory even though I am uncertain how this would be accomplished by Rawls. One concern I have is that Hampton's argument relies on her employment of a legal definition of actual contracts. For instance, Hampton appeals to the voluntaristic character of real contracts by claiming that contractual agreements are not necessarily irrevocable because all parties may agree to cancel the contract. Just as making a contract is voluntary so too is ending it as long as all parties agree. Hampton believes that whether the original position can be cashed out in contractarian terms depends on whether the finality condition (i.e., that the principles of justice are the last courts of appeal when distributing resources or arbitrating conflicts) can be replaced by the irrevocability condition (i.e., once the veil of ignorance has been lifted the principles chosen are final). But by employing the legal definition of contracts it is argued that the irrevocability condition is too weak to accomplish this because the parties may reason that if all agree, there is a chance that the original contract may be reversed once the veil is lifted. Hence, the finality condition would not hold. I have some reservations about Hampton's characterization of this possibility. She claims that once individuals discover that the consequences of the two principles of justice are distasteful then they can go back and renegoiate the contract. But her description seems to imply that when renegotiating the contract we do so with full knowledge. This is not so for Rawls. When renegoiating the contract we re-enter the original position and hence, we will likely end up with the same principles. Or perhaps, alternatively, the general conception of justice will be appealed to in these circumstance. Rawls does not seem committed to accepting a legal definition of contracts because applying such a notion presupposes the existence of institutions. Still, he must provide some justification for introducing formal conditions restricting the two principles. Nonetheless, I agree with the spirit of Hampton's criticisms and more importantly, if Rawls cannot describe the original position in contract terms perhaps he should have presented it as a reasoning process simply.

The initial situation provides the conditions necessary to ensure that the arguments for or against particular principles of justice are not grounded on considerations that are unfair. The principles that are arrived at are the principles that "...free and rational persons concerned to further their own interests would accept in an initial position of equality as defining the fundamental terms of their association."4 The initial situation of agreement must be described in such a way that the principles that result are both fair and reasonable. Thus, the veil of ignorance is introduced as a representation of the kinds of constraints that can be fairly imposed on the rational choice for principles of justice.

The veil of ignorance also guarantees pure procedural justice. Pure procedural justice "...obtains when there is no independent criterion for the right result: instead there is a correct or fair procedure such that the outcome is likewise correct or fair, whatever it is, provided that the procedure has been properly followed."5 The veil of ignorance represents the necessary constraints by restricting the knowledge of facts of one's particular circumstances and one's attributes in order to nullify "...the effects of specific contingencies which put men at odds and tempt them to exploit social and natural circumstances to their own advantage."6 If the veil of

---

4John Rawls, A Theory of Justice, (Cambridge: Belknap Press), p. 11.
5A Theory of Justice, p. 86.
6Rawls, p. 136.

ignorance is not in place, individuals can tailor the principles of justice to their advantage. And doing so is contrary both to justice as fairness and the realization of pure procedural justice.

When the veil is in place the hypothetical people of the initial situation do not know their particular class or social status in society. They also lack knowledge of their natural assets and abilities, conceptions of the good, their rational life plans and the special features of their own psychology. It is also assumed that they lack knowledge of their society such as its political, cultural, and economic development. Nor do they know their historical placement. However, they are rational, they know that the circumstances of justice obtain and they also know that they may need the primary social goods to realize their conceptions of the good. The parties also have access to general information necessary for choosing principles of justice. For example, they have information on general laws, and theories, political affairs, economic theory, social organization, and the laws of human psychology.7 If the participants of the original position lack all knowledge or any ties to the empirical world it is doubtful that determinate principles can be chosen at all. The exclusion of knowledge of our particular circumstances ensures that the principles of justice are not prejudiced by our individual conceptions of the good and that the principles accurately reflect the

---

7Rawls, p. 137.

sense of justice that is common to all citizens in society. An important feature of justice as fairness is that

> ...to each according to his threat advantage is not a principle of justice. If the original position is to yield agreements that are just, the parties must be fairly situated and treated equally as moral persons. The arbitrariness of the world must be corrected for by adjusting the circumstances of the initial contractual situation.8

This naturally leads us to a consideration of the principles themselves. The liberty principle and the difference principle are described as follows:

> liberty principle: Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all.
>
> difference principle: Social and economic inequalities
> are to be arranged so that they are both:
> (a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and
> (b) attached to offices and positions open to all under conditions of fair equality of opportunity.9

They are lexically ordered to ensure that liberty is not sacrificed for economic advantage and therefore "...liberty can be restricted only for the sake of liberty."10 The difference principle represents Rawls' view of distributive justice which is motivated by his belief that one's assets and talents should be treated as if they are common assets.

Section 1. The Primacy of Justice

Sandel misinterprets what the primacy of justice means for Rawls in two ways. First Sandel believes that if

---

8Rawls, p. 141.
9Rawls, p. 302.
10Ibid.

justice is epistemologically prior to conceptions of the good, then the principles of justice must be derived independently of all societal values. Sandel then argues that this entails a metaphysical conception of the self that is prior to and independent of its attributes.11 His argument is unsound because the two principles need not be independent of all societal values. More importantly, if Rawls' project as a whole is taken into consideration the principles of justice must include at least some judgements about value.

There is a similar defect in Sandel's argument whereby he derives the transcendental self for Rawls' view of the priority of procedure. Briefly, Sandel claims that for Rawls the justification for the procedure must come antecedent to the justification of the principles themselves. In other words, the procedure (i.e., the description of the original position) must be justified prior to the principles that it is designed to produce. Just as the subject reflects the priority of justice, it is argued that the subject also must reflect the priority of procedure. Hence, Sandel contends that Rawls' view of the primacy of justice and procedure are founded on the metaphysical conception of the self that is prior to and independently identifiable from its ends. In the latter case, I argue that the justification for the original position is not strictly independent of all our values as

11In chapter II I argue that this conceptions must be viewed as metaphysical.

interpreted by Sandel. This is supported by Rawls' coherence theory of justification which is embedded in the notion of reflective equilibrium. This shows that the justification for the procedure is in part dependent on the principles of justice and our considered convictions about justice.

Lastly, Sandel argues that if justice is primary other virtues that are equally worthy will be eliminated. But, Sandel falsely assumes that in order for justice to be primary, actual conflict and the application of the two principles must be necessary conditions.

Section A. Justice as a First Virtue

Sandel's most sweeping claim is that deontological liberalism is flawed because it demands that justice is primary; it demands that justice be the first virtue of society. As a communitarian, Sandel's conception of the ideal society is considerably different. In this society the principles of justice are absent in some circumstances because they are unnecessary. The virtues of benevolence and fraternity replace the need for principles of justice by eliminating conflict in particular situations and by reducing conflict in society generally. For Sandel, this view of society is morally superior because it need not invoke justice to resolve all conflicts whereas deontological liberalism he believes, must presuppose both conflict and the need for the principles of justice. Sandel argues that

> ...if the virtue of justice is measured by morally diminished conditions that are its prerequisite, then the absence of these conditions however this state of affairs might be described-must embody a rival virtue of a least commensurate priority, the one that is engaged insofar as justice is not engaged.12

The objective and subjective circumstances of justice (i.e., the moderate scarcity of resources and different interests and conceptions of the good life) provide the background for social justice. Consequently, Sandel claims that if justice is the first virtue of social institutions and the circumstances of justice obtain, this represents a less desirable situation from a moral standpoint, than if there is no need for principles of justice. In society in which the members are strongly aligned with the community's values and goals considerations of justice rarely arise. Sandel maintains that there are a variety of human associations that exemplify both an allegiance to community values and the relative absence of the circumstances of justice. For example,

> These would include, at various points along the spectrum, tribes, neighborhoods, cities, towns, universities, trade unions,.... (and) a wide variety of ethnic,religious, cultural, and linguistic communities with more or less clearly-defined common identities and shared purposes, precisely those attributes whose presence signifies the relative absence of the circumstances of justice. Although the circumstances of justice might well exist in all of these cases, they would not likely predominate, at least to the extent that justice was engaged in all cases in greater measure than any other virtue.13

12Sandel, pp. 31-32.
13Sandel, p. 31.

Insofar as the circumstances of justice do not obtain and conflicts are scarce the climate of these associations must be considered a "...virtue of at least correlative status."14 In these situations he maintains that where there is little conflict the principles of justice rarely apply and hence, justice is not in every case the first virtue. Furthermore, he suggests that in these existing associations justice is not prior to other virtues such as benevolence in all cases and hence, social justice is not likely to be an issue. In Sandel's opinion, the presence of conflict betokens the inferiority of Rawlsian society. However, I argue that on a more plausible interpretation of the primacy of justice this criticism may be answered.

Even though Sandel concedes that Rawls' account can allow that benevolence is a virtue, Sandel's real complaint is that A Theory of Justice cannot allow communitarian concerns to outweigh inconsistent individual concerns.15 On Sandel's account the "correct" conception of the good for any individual in society is promotion of communitarian ends. In opposition, Rawls allows for all conceptions of the good as long as they are compatible with the two principles of justice and do not block the satisfaction of equally legitimate conceptions of the good. Perhaps only as a procedural constraint, Rawls rejects the communitarian view of society when he states that "...we do not want to rely on an undefined conception of community, or to suppose that

14Sandel, p. 32.
15Sandel, pp., 61-63.

society is an organic whole with a life of its own distinct from and superior to that of all its members in their relations with one another."16 But for Sandel and many communitarians society is viewed as superior to its members in just this way.

Sandel's main claim is that "...if justice depends for its virtue on certain empirical preconditions, it is unclear how its priority could unconditionally be affirmed."17 However, it can be argued that justice remains the first virtue because conflict and the application of the principles can be viewed as sufficient but not necessary conditions for the primacy of justice. One might be initially sympathetic to Sandel's interpretation of Rawls on the basis of these and similar passages. "Justice is the first virtue of social institutions, as truth is to systems of thought."18 "Unless these circumstances existed there would be no occasion for the virtue of justice,...".19 However, the first passage does not commit Rawls to the view that, for justice to be primary, conflict and the application of the two principles are required as prerequisites. And the second passage does not preclude the possibility that justice is the first virtue even if justice is guaranteed. Being the first virtue, justice need not be appealed to if it is already guaranteed. In this situation,

---

16Rawls, p. 264.
17Sandel, p. 30.
18Rawls, p. 3.
19Ibid, p. 128.

the principles of justice are not exercised because there is no conflict to be resolved.

For instance, consider Rawls' claim that "a theory however elegant and economical must be rejected or revised if it is untrue; likewise laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are unjust."20 Justice and truth cannot be sacrificed for other virtues where they are not guaranteed. Thus, truth and justice are the most important considerations when either choosing or constructing a system of thought or a theory of justice. Rawls informs us that "being first virtues of human activity, truth and justice are uncompromising."21 But to claim that other considerations cannot override justice for instance, does not require that conflict or the application of the principles are necessary conditions in situations where justice is already the status quo. Hence, the status of justice as the first virtue remains unaffected. In the case of justice, conflict and the application of these principles may be evidence merely for the fact that justice is not guaranteed by the type of association itself. If conflict and the application of the principles are viewed as sufficient conditions Rawls need not deny that where the circumstances of justice are absent that the two principles of justice apply. Moreover, he is not committed to the claim that where justice is guaranteed it is not the first

---

20Ibid, p. 3.
21Rawls, p.4.

virtue of institutions. In a situation of social cooperation principles may remain on the shelf, so to speak, as insurance that if a conflict arises involving questions of justice for social institutions they may be used in appropriate circumstances. This is clearly Rawls' view for he states that,

> ...justice is the virtue of practices where there are competing interests and where persons feel entitled to press their rights on each other. In an association of saints agreeing on a common ideal if such a community could exist, disputes about justice would not occur.22

It is plausible to assume that in an association of saints, the lack of disputes about justice implies that justice is already being met.23 For Rawls, the principles of justice apply in situations where competing interests escalate in to actual conflicts about justice. They obtain "...whenever mutually disinterested persons put forward conflicting claims to the division of social advantages under conditions of moderate scarcity."24 However, despite the satisfaction of the circumstances of justice, justice remains the first virtue. In short, Rawls believes that the circumstances of justice are sufficient but not necessary conditions in order for justice to serve as the first virtue. Yet, Sandel believes that the list of associations he cites provides a counter-example to such a claim because even though the circumstances obtain they are prevalent to a relatively

---

22Rawls, p. 129.
23I suspect that for Rawls, an association of saints where questions of justice do not arise, may be possible for some beings but not for human beings.
24Rawls, p.128.

small degree. Justice as a virtue applies to a lesser extent in some cases than other virtues. Regardless of this, for Rawls justice as the first virtue is not jeopardized because he need not view the circumstances as necessary conditions.

Sandel states that we can imagine a range of intimate associations where the circumstances of justice are not prevalent. Further, in reference to Hume, he says that "we need not have recourse to utopian visions or the fictions of poets to imagine such conditions, but 'may discover the same truth by common experience and observation'".25 The previous claim suggests an ambiguity in how we arrive at these conditions. If all that is required is to imagine an association where the circumstances of justice do not obtain this is possible but it does not serve as a counter-example. In short, imagining situations where these conditions do not obtain does not make this so empirically. Perhaps Sandel's believes that the existence of these conditions are empirical facts. His reference to Hume seems to suggest this interpretation. However, this cannot be the case because in these situations the circumstances do obtain, albeit only to a lesser degree. This reaffirms Rawls' claim that the circumstances of justice are features of existing societies.26 For Rawls, these principles must be chosen because of the peculiarity of societies in general. Rawls assumes that "... society is a more or less self-sufficient

25Sandel, p. 31.
26Rawls, p. 129-130.

association of persons who in their relations to one another recognize certain rules of conduct as binding and who for the most part act in accordance with them."27 Even though individuals in society cooperate socially they have different interests which also may lead to disagreements about the distribution of social goods thereby generating conflicts. Rationally, all prefer a greater share of social goods so they can more fully realize their goals.28 Consequently, the principles of justice provide "...a way of assigning rights and duties in the basic institutions of society and they define the appropriate distribution of the benefits and burdens of social cooperation."29

The need for the principles arises from certain background conditions. These conditions are the objective and the subjective circumstances of justice. One of the most important of the objective conditions is moderate scarcity. Moderate scarcity is described as a situation in which "natural and other resources are not so abundant that schemes of cooperation become superfluous nor are conditions so harsh that fruitful ventures must inevitably breakdown."30 In short, moderate scarcity initiates cooperation but cooperation need not eliminate moderate scarcity. Hence, the important question for Rawls is "How should goods be distributed?" While the objective circumstances of justice are concerned with the external

---

27Ibid., p.4.
28Ibid.
29Ibid.
30Rawls, p. 127.

forces that may thwart the realization of our goals, the subjective circumstances of justice are primarily a consideration of the general characteristics of individuals within society. The significant aspect of the latter, is that although individuals have similar needs and interests they also have individual plans of life that may reflect different conceptions of the good. "These plans, or conceptions of the good, lead them to have different ends and purposes, and to make conflicting claims on social and natural resources available."31 There are many levels of similarities and differences between individuals that may either lead to conflicts or reduce conflicts. For instance, if we all want the same goods, and resources are moderately scarce, this may lead to conflicts about distributive justice. Yet, if we all want different goods this may not only make social cooperation possible but also may reduce the occurance of conflicts. On a second level we may have either similar or different conceptions of the good life. If we happen to have the same values this does not necessarily eliminate conflicts because there may be disagreements about the best means to achieve those values. Moreover, it is clear that having different conceptions of the good can lead to conflict if individuals promote those conceptions at a public level. Conflicts may arise where there are disagreements about which conception of the good should be encouraged. Hence, for Rawls justice remains a

---

31Ibid.

problem for society because at the first level we do have similar needs and at the second level we do not all share the same conception of the good. And even if we did agree on a conception of the good this does not entail a consensus on the best means for achieving the good.

The circumstances of justice provide the necessary foundation for mutually beneficial cooperation, but they are also the conditions that serve as the backdrop for issues involving the resolution of conflicts concerning social justice. The circumstances of justice are generalizations from empirical facts about society. Given the nature of human associations questions of justice cannot be avoided because

> ...a human society is characterized by the circumstances of justice. The account of these conditions involves no particular theory of human motivation. Rather, its aim is to include in the description of the original position the relations of individuals to one another which set the stage for questions of justice.32

It may be the case that in different societies the background conditions for justice exist in different degrees. But, it is plausible to assume that in every actual human society the circumstances of justice obtain to the degree sufficient to make justice a problem.

Sandel claims that if individuals within a group have similar values and ends justice prevails to a relatively small degree, reasoning that justice is not engaged in greater measure than any other virtue. But I explain that

---

32Rawls, pp. 129-130.

if we have similar "ends" in the sense of having similar "needs" then at the first level this may be a source of conflict if resources are scarce. On the second level having the same conception of the good need not eliminate conflict because individuals may disagree on the means of achieving those ends. Rawls clearly agrees with my interpretation of the nature of conflict because he states that "persons generally want similar sorts of things, liberty, opportunity, shelter and nourishment, yet these wants may put them at odds."33 Moreover, even if the members of a group have similar values or goals justice may still be necessary to regulate the way these goals are realized. The problem is that by merely eliminating the subjective circumstances of justice and not also eliminating moderate scarcity there is no compelling reason to suppose that all conflicts will disappear. Perhaps Sandel believes that if our values and goals coincide closely we will also agree on the way that these are to be achieved. But, if this is his claim it is doubtful that there is empirical support for it. In any event, it is quite clear that common purposes do not entail consensus and in fact these common purposes may at times initiate conflict if the objective circumstances obtain. This is confirmed by Charles Taylor in "Interpretation and the Sciences of Man". In this article, shared purposes and meanings of a community are common aspirations that hold as reference points to all

---

33Rawls, p.526.

debate about such a common goal. They are "rooted in social practise."[34] But communities that have shared purposes and meanings are not necessarily communities that lack conflict. In fact Taylor believes that common meanings and purposes are differently articulated by different groups and are "...very often the cause of the most bitter lack of consensus."[35] Taylor and Sandel share many views. Yet, this is a case where Sandel fails to consider the complex nature of conflict.

We have yet to examine the status of the circumstances of justice for Rawls. The objective condition of justice is clearly an empirical generalization about the world. We do find ourselves in conditions of moderate scarcity. Mutual disinterest (i.e., the condition of not being interested in someone else's interests) reflects the empirical fact that even though we may have similar interests and needs we also have our own plans of life which often leads us to have different purposes and ends. This in turn, may initiate conflicting claims for social goods. In other words, it is an empirical fact that individuals have different interests that may serve as a source of conflict. But on another level mutual disinterest is procedural. We do not want to assume as part of the framework of the theory that people are generally mutually disinterested, for this is surely not the case; but rather, assuming this is less demanding and

---

34Charles Taylor, "Interpretations and the Sciences of Man," in Philosophical Papers 2, p. 37.
35Taylor, p. 39.

weaker than assuming that all people are either selfish or benevolent. This is a weaker assumption because it avoids controversial issues about human nature. Rawls is non-commital about this issue because it has not been resolved and hence he does not believe that it can be included as part of the structure of his theory. This allows Rawls to avoid making false claims that all people are motivated by either selfish or benevolent interests. Furthermore, it reflects the considered judgement that conceptions of the good are morally irrelevant. Hence, by assuming neither we avoid assuming that people are one or the other. Motivational factors may be determined by the uncontested general facts and theories that the parties in the original position are taken to know. Mutual disinterest does not presuppose either of these views. Assuming mutual disinterest takes into account that people may be selfish or that they may be benevolent.

Sandel believes that certain relationships such as friendship and associations like the family are often governed by spontaneous affection. In addition, he objects to justice as the first virtue on the grounds that other virtues such as affection, fraternity and benevolence are replaced by an overwhelming concern for justice.[36] In the following Sandel describes a scenario in which his concern might be realized:

> If, out of a misplaced sense of justice a close friend of long-standing repeatedly insists on

---

36Sandel, p. 33.

> calculating and paying his precise share of every common expenditure, or refuses to accept any favor or hospitality except at the greatest protest and embarrassment, not only will I feel compelled to be reciprocally scrupulous but at some point may begin to wonder whether I have not misunderstood our relationship.37

Justice in this situation is inappropriately applied and comes at the cost of eliminating other virtues that make justice unnecessary. This example is inapplicable to Rawls' theory, however, because the principles of justice are only applicable to the basic structure of society where in general these relationships play no role. In other words,

> these principles may not work for the rules and practises of private associations or for those of less comprehensive social groups. They may be irrelevant for the various informal conventions and customs of everyday life;....38

Hence, in most cases this kind of example is inapplicable to the basic structure of a just society. But let us consider a case in which friendship and family relations do make contact with the basic structure of society. First, let us assume that we live in a just society. Second, let us suppose that my son or my closest friend murders a police officer. I might ask myself whether I should protect him from the law or turn him in. I believe that most people agree that the just thing to do is to turn him in to the police and accept the fact that my son or friend is a murderer. In this case, the two principles may be appropriately applied because it involves an issue of social justice. But, in the case that Sandel discusses the two

---

37Sandel, p. 35.
38Rawls, p. 8.

principles of justice are inapplicable because there is no issue of social justice to be resolved. But I doubt that Sandel disagrees with me here because the character of the family (e.g., common purposes, affection, and feelings of fraternity) that often govern the relations within the family or between friends is in this case inappropriate and comes at the cost of social justice.

Section B. Justice and Procedure

In "Justice as Fairness: Political not metaphysical" Rawls claims but does not justify the proposition that "...Michael Sandel is mistaken in supposing that the original position involves a conception of the self '...shorn of all its contingently-given attributes', a self that 'assumes a kind of supra-empirical status,...' and given prior to its ends, a pure agent of agency and possession, ultimately thin."39 This is the proposition that I shall defend.

Sandel extracts three propositions from A Theory of Justice to construct his versions of the Rawlsian selves. The first is that justice and the right are prior to any conception of the good. The second, is that procedure is prior to the principles that it justifies. The third, is that our talents and assets are not deserved and hence, are morally arbitrary. Sandel ascribes two conceptions of the self to Rawls. Initially, these conceptions are viewed from

---

39John Rawls, "Justice as Fairness: Political not metaphysical" in Philosophy and Public Affairs, Summer 1985,vol. 14, no.3, p. 239.

a normative point of view as part of a moral theory. Hence one expects their descriptions to be couched primarily in ethical terms reflecting a particular view of human agency and attributes relevant to a discussion of such a theory. Nonetheless, the descriptions of these conceptions are most often metaphysical. In his discussion of the Kantian ego, Sandel makes a distinction between the metaphysical description and the moral description of the subject. I argue in this case, that the entailment that Sandel requires between the moral and the metaphysical subject cannot be made. Much of Sandel's project depends on the entailment thus described. In the case of the constitutive conception of the self the distinction between the moral and metaphysical features of the self becomes blurred. Hence, in the latter case I also argue that Sandel conflates ethics with metaphysics.

Before I examine whether the Kantian self follows from the priority of justice I explain what this means for Rawls. His theory is one in which the principles of justice and right are both morally and epistemologically prior to conceptions of the good; that is to say, the principles of justice and right cannot be outweighed by individuals implementing conceptions of the good nor can they be derived as principles by appealing to conceptions of the good. Hence, for Rawls,

> Each person possesses an inviolability founded on justice that even the welfare of society as a whole cannot override. For this reason justice denies that the loss of freedom for some is made right by

> a greater good shared by others....Therefore in a just society the liberties of equal citizenship are taken as settled;the rights secured by justice are not subject to political bargaining or to the calculus of social interests.40

Even though individual conceptions of the good cannot outweigh the principles of justice and the right, this does not imply that conceptions of the good are unimportant or unnecessary. The importance of individual conceptions of the good is compatible with the principles of justice by disallowing them from entering into the discussion whereby the principles of justice will be chosen.41 Yet, some view of the good must provide the material for arriving at determinate principles. For instance, Rawls believes that if the people participating in the hypothetical original position lacked all ties to empirical conditions such as at least some conception of the good that they would be "...lacking explicable connections with human conduct...."42. The "thin theory of the good" provides one such connection to the empirical world. It assumes that those choosing principles want primary social goods such as self-esteem, liberty, opportunity, wealth, and income.43 This conception of the good (the thin theory of the good) ensures, Rawls thinks, that determinate principles can be chosen. But one might wonder whether this violates the restriction that the two principles must be derived

---

40A Theory of Justice, pp. 3-4.
41I discuss this in more detail in the last section of the final chapter.
42A Theory of Justice, p. 256.
43A Theory of Justice, p. 396.

independently of the conceptions of the good. Rawls reveals the benign nature of the thin theory of the good when he states that "...wanting (primary social goods) is part of being rational; and while each is presumed to have some conception of the good, nothing is known about his final ends."44 The notion of primary social goods, Rawls contends, is limited so that it cannot be viewed as influencing one's choice of principles on the basis of a full conception of the good. A deontological theory is distinguished from a teleological theory on the grounds that in the former, the principles of justice and the right are derived independently of conceptions of the good. The latter, however, reduce the right to terms of whatever they view to be the good.45 Moreover, in deontological theories the good cannot outweigh the right whereas teleological theories having reduced the right to terms of what they view to be the good, the right is not similarly secured.46 Hence, being the first virtue of social institutions, justice for Rawls is both morally and epistemologically prior to the good because in the former when the two conflict, the principles of justice outweigh the good and in the latter they are derived independently of our complete conceptions of the good.47 Moreover, even though justice is

---

44A Theory of Justice, p. 253.
48A Theory of Justice, p. 30.
46A Theory of Justice, p. 28.
47A Theory of Justice, p. 396. Our complete conceptions of the good are what we attain after the veil of ignorance is lifted and we have full knowledge of our circumstances, rational plans of life and assets.

part of a conception of the good I later argue in a similar fashion that this does not jeopradize the deontological nature of his theory because it is not a full conception of the good nor is it arbitrarily determined.

Sandel contends that the justification for the epistemological priority of the two principles is based on the need to derive regulative principles to appraise the standards of society. He argues that if principles of justice and the right are themselves based on existing societal values, they cannot serve as principles for evaluating institutions. Sandel's interpretation of the primacy of justice and its relation to the self is put simply when he states that "as the priority of justice arose from the need to distinguish the standard of appraisal from society being appraised, the priority of the self arises from the parallel need to distinguish the subject from its situation."[48] To introduce the first of his three arguments for the Kantian self Sandel focuses on the essential feature of deontological liberalism. Since the principles of justice and the right are derived independently of conceptions of the good, justice is primary. Sandel is convinced that if justice is prior then this entails a subject that is prior to and independent of its ends, aims and wants.[49] He interprets the priority of the deontological subject in two different ways. First, since the principles of justice are morally prior to conceptions

---

48Sandel, p, 20.
49Sandel, pp. 19-20.

of the good (i.e., considerations of justice outweigh considerations of the good) the subject must also be morally prior to its ends, aims, and wants. Priority in this sense reflects the "...imperative to respect above all the autonomy of the individuals, to regard the human person as the bearer of a dignity beyond the roles that he inhabits and the ends he may pursue."[50] In short, we are required to view people as choosers. The second sense of priority is derived from the epistemological priority of the principles of justice over conceptions of the good. Hence, Sandel asserts that if the principles of justice must be derived independently of conceptions of the good and existing societal values the subject must be "...prior to the ends it affirms - prior in the sense of independently identifiable" and this Sandel claims, "is an epistemological requirement."[51] Sandel describes this sense of priority as epistemological, yet it is a metaphysical claim about the self. Since this conception of the self is one in which the subject must be antecedent to any empirical conditions it is termed the dispossessed, disembodied or Kantian conception of the self,[52] wherein the moral subject conforms to a "voluntarist" view of human agency.[53] Hence, via the epistemological view of justice we arrive at the Kantian or

---

50Ibid., p.20.
51Sandel, p. 20.
52Sandel, p. 13.
53The voluntarist view of human agency is pitted against Sandel's own "cognitive" view of human agency. In the former, individuals actively choose their ends whereas in the latter, individuals discover their ends.

transcendental self. Correspondingly, via the moral priority of justice we arrive at a voluntarist account of human agency.

It is true that Rawls claims that the "...self is prior to the ends which are affirmed by it;...".54 Yet, this suggests that the self is prior from a moral point of view rather than prior from a metaphysical point of view. For Rawls further says, "...even a dominant end must be chosen among numerous possibilities."55 Hence, for Rawls the subject is prior to its ends in the sense that, since ends are not merely given, they must be chosen in ways which mirror the subject Sandel derives from the moral priority of justice. There is insufficient evidence from Rawls' claim about the subject to suggest that the subject must be epistemologically prior to its ends. A reconstruction of Sandel's argument will reveal his mistake.

1. Justice and the right are both morally and epistemologically prior to conceptions of the good.

2. The epistemological priority of the principles of justice is represented by deriving the principles independently of conceptions of the good.

3. If the principles of justice are based on existing societal values they cannot serve as regulatory principles for evaluating the institutions of society.

4. Hence, the subject must reflect the epistemological priority of justice over the good.

5. Therefore, the subject must be distinguished from its situation and must be identifiable independently from its ends.

54A Theory of Justice, p. 560.
55Ibid.

Premises 1. and 2. reflect an accurate characterization of Rawls' account. Sandel's error lies, in part, in premise 3. This propositions states that the principles of justice must not be grounded in any existing societal values, otherwise they could not provide an independent standard for assessing our institutions. Sandel infers from this premise that the principles cannot include an account of any societal values if their regulatory capacity is to be upheld. It is true for Rawls' theory as a deontological one that not all societal values can be assumed, hence particular full conceptions of the good have no place when arguing for the principles of justice. But it does not follow, in turn, that no societal values can be assumed. The characterization of Rawls' theory as one part of moral theory in general, requires that our considered convictions about justice form part of the basis for either accepting or rejecting the two principles. "Justice as fairness is a theory of our moral sentiments as manifested by our considered judgments in reflective equilibrium."56

In Rawls' discussion on moral theory, he explains that it is an attempt to describe our moral capacity. Having a moral capacity is a skill in judging a thing to be right or wrong in normal circumstances and justifying that judgment by reasons. Moral theory is an attempt to formulate a set of principles that reflect this capacity.57 A Theory of Justice as political theory is one aspect of moral theory

56A Theory of Justice, p. 120.
57A Theory of Justice, p.46.

and describes our capacity for "...judging things to be just or unjust, and in supporting these judgements by reasons."[58] It is an attempt to formulate a set of principles that "...when conjoined to beliefs and knowledge of the circumstances, would lead us to make these judgements with their supporting reasons were we to apply these principles conscientiously and intelligently."[59] Hence, the principles are designed to articulate our sense of justice that upon reflection is accepted. Therefore, they must express at least some of our social values and considered convictions about justice. Hence, even though full conceptions of the good are ruled out views about justice are not.

In an effort to clarify the role of moral theory Rawls compares it to linguistics. For instance, our sense of grammaticalness or our ability to recognize well-formed sentences is "...expressed by principles which make the same discriminations as the native speaker."[60] Rawls, believes that just as our grammatical principles articulate our sense of grammaticalness, the principles of justice articulate the best account of our sense of justice once considered in reflective equilibrium.

It might be asked, "If our principles of justice are not independent of certain judgements about justice, how can these principles provide an independent standard for evaluating our institutions in society?" The answer to this

---

58Ibid.
59Ibid.
60A Theory of Justice, p. 47.

question can be found in Rawls' discussion of reflective equilibrium.

Reflective equilibrium is part of Rawls' epistemological framework which functions as a means of justifying the description of the original position, the principles of justice, and our considered convictions about justice. One way of justifying the description of the original position is to see whether the principles of justice that are chosen "...match our considered convictions of justice or extend them in an acceptable way."[61]

First, we start by characterizing the description of the original position based on "generally shared and preferably weak conditions."[62] When we arrive at the point that determinate principles can be chosen we then see whether the principles of justice we arrive at are consistent and cohere with our considered convictions about justice. When there is a discrepancy between the two we can alter the description of the original position or we can revise our judgements.[63] We eventually

> find a description of the initial situation that both expresses reasonable conditions and yields principles which match our considered judgements duly pruned and adjusted. This state of affairs I refer to as reflective equilibrium because at last our principles and judgements coincide; and it is reflective since we know to what principles our judgements conform and the premises of their derivation.[64]

---

61 A Theory of Justice, p. 19.
62 A Theory of Justice, p.20.
63 Ibid.
64 Ibid.

We must have principles that articulate our considered judgements that have been accepted upon reflective consideration.

Our considered judgements or convictions may be intuitive judgements but they are ones in which we have the "...greatest confidence..."65, and which we may not be willing to reject. "For example, we are confident that religious intolerance and racial discrimination are unjust."66 Rawls does though allow that there are judgements which are erroneous and may be rejected.67 He believes that the principles of justice reflect our undistorted and accepted considered convictions about justice.

The best account of this is not merely a matter of seeing whether the principles match our judgements unreflectively but "...rather the one that matches (our) judgements in reflective equilibrium..., (T)his state is one reached after a person has weighed various proposed conceptions and he has either revised his judgements to accord with one of them or held fast to his initial convictions (and corresponding conception)."68 Reflective equilibrium is a balance between a conception of justice and one's judgement after the reasons for making that judgement

---

65A Theory of Justice, p. 19.
66Ibid.
67On page 48 of A Theory of Justice, Rawls claims that erroneous judgements are often based on hasty decisions, uncertainty or made in unfavorable emotional circumstances. However, it may turn out that even under favorable circumstances some of our judgements will be rejected.
68A Theory of Justice, p. 48.

have been explained and examined. This may result in either revising, rejecting or accepting certain of those judgements. It is a process of going back and forth between our convictions about justice and various conceptions until we have reached an equilibrium or balance between our convictions and the best account of them once reflected upon, and this account forms the principles of justice that articulate our sense of justice. In general, the justification for A Theory of Justice is determined by how the conception "...fits in with and organizes our considered judgements in reflective equilibrium.... Justification is a matter of the mutual support of many considerations of everything fitting together into one coherent view."69

For Rawls the principles of justice include some societal values because they reflect our sense of justice. However, "justice is not to be confused with an all-inclusive vision of the good society, or thought of as identical with the conception of right." But rather "it is only one part of any such conception and it is but one species of right".70 At this point we might wonder whether Rawls' theory is truly a deontological theory if justice is part of a conception of the good. There are at least two reasons why his theory remains deontological in character. First, justice is a species of the right and although it is part of a conception of the good it excludes full

---

69A Theory of Justice, p. 579.
70John Rawls, "Justice as Reciprocity," in Choice and Action, ed., by Charles Reid, (New York: Macmillan Publishing Company), p. 293.

conceptions of the good. Hence, it is determined independently of the good as the satisfaction of rational desire, pleasure or other views of the good. By excluding knowledge of our assets and circumstances as reasons for accepting a conception of justice we have retained the deontological nature of the theory. By so doing justice need not maximize views of the good. Second, even though Rawls' conception of justice includes some societal values about justice it is a mistake to assume that the acceptance of certain values over others is merely arbitrary. If the acceptance of judgements about justice is arbitrary then Rawls' theory as a deontological one may be problematic because there may not be a way of distinguishing between judgements that represent views that are derived from full conceptions of the good that are not shared by all and those judgements that could form a shared basis for a conception of justice. But, in arriving at the principles of justice we also assess our existing societal values and principles that determine the distribution of social goods. The justification for our judgements about justice must rest on sound and reasonable arguments. Hence, those judgements that cannot be adequately defended will either be rejected or revised. Only those that can be justified are accepted as comprising our considered convictions about justice. In reflective equilibrium we go back and forth between our considered convictions about justice and various conceptions of justice until we have reached a balance between our

convictions and the best account of them once reflected upon. At this point we can feel confident that the principles of justice reflect our sense of justice.

Rawls' argument against justice as involving desert illustrates one case of how we can distinguish judgements about justice that we think can provide a shared basis for a conception of justice and those judgements that upon reflection are not part of our fixed moral judgements. Our natural assets and social positions are not solely the result of our own actions, and hence we can make no claims to deserve the benefits that result from these. Our characters and our placement in the social structure are the products of the natural lottery. Furthermore, the

> ...assertion that a man deserves the superior character that enables him to make the effort to cultivate his abilities is equally problematic; for his character depends in large part upon fortunate family and social circumstances for which he can claim no credit. The notion of desert seems not to apply to these cases.71

Rawls' argument against desert reflects the view that society influences and shapes the characters that we have and even the goals that we pursue. We are not responsible for bringing about all of our features. Yet, he also believes that we are capable of individual control. In short, Rawls believes that even though society has some influence over us this is compatible with choice. But this latter claim does not then imply that there are some things that we deserve in a pre-institutional sense. Hence, if the

---

71A Theory of Justice, p. 104.

argument about desert is successful the constraints in the original position should be acceptable.

Rawls argues that the principle of desert cannot be chosen as a distributive principle in the original position because once reflected upon it is shown that "...it is one of the fixed points of our moral judgments that no one deserves his place in the distribution of natural assets any more than he deserves his initial starting place in society."72 If Rawls' argument defending the moral arbitrariness of our assets is sound, then upon reflection we reject desert as a distributive principle for the basic structure in cases of justice. We are shown that our attraction to desert as a distributive principle is in fact contrary to what we really believe is just. If one of the functions of A Theory of justice is to question our existing societal values then desert as an appropriate distributive principle must also be questioned. Hence, Rawls' argument rejecting a pre-institutional notion of desert illustrates the reasoning process whereby we come to accept,reject, or revise our existing judgments about justice. Moreover, reflective equilibrium is a process whereby we assess the legitimacy of our existing values and judgments about justice. The rejection of desert is in part determined by the fact that it is contrary to what we view to be just upon reflection. Reflective equilibrium serves as a complex but rather ingenious method for isolating from the two

---

72A Theory of Justice, p. 311.

principles of justice our full conceptions of the good and our judgements about justice that do not form part of our fixed moral judgements. The principles of justice are derived independently of arbitrary judgments about justice and full conceptions of the good life. Thus, the principles of justice as regulatory principles also provide a standard independent of existing societal values for evaluating institutions insofar as reflective equilibrium filters out judgments about justice that may not be publicly accepted and hence, cannot provide a shared basis for a conception of justice. In the case of desert, it is rejected also because its very existence depends on the existence of institutions. And, to include it as a distributive principle may be to violate the regulative role of the two principles and the deontological nature of the theory by allowing the fact that we have certain assets, and social positions, and conceptions of the good to slip in the back door. In effect, if desert is included as a first principle of distribution this may be merely to rubber stamp the way institutions already distribute social goods. But, whether this is just or not is the very question that Rawls poses. Therefore, the principles of justice not only articulate our sense of justice but they do so without presupposing that any one of our judgments about justice is sound or presupposing a full conception of the good thereby, ensuring the deontological nature of the theory.

With this in mind, it is clear that Sandel misinterprets Rawls' view of the epistemological priority of justice. For Rawls, if the principles of justice are to reflect our sense of justice they may be based on some societal values. But they must do so in such a way as not to include values that must be rejected or revised. Sandel's conclusion at 5 does not follow because the success of his argument depends on all the premises being true and premise 3 is clearly not true for Rawls. Therefore, the Kantian self does not follow from the epistemological priority of justice.

Sandel makes a similar error in his discussion as to whether taking the original position as a procedural rendering of Kant divorces A Theory of Justice from dubious Kantian metaphysics. Sandel explains that for Rawls the original position is to be viewed as a hypothetical contract (i.e., it describes a procedure) which embodies the Kantian ideas of autonomy and the categorical imperative.[73] By restricting our knowledge of assets and circumstances the veil of ignorance expresses our freedom from contingencies that are morally arbitrary. The idea of a categorical imperative is represented by making the principles of justice the object of rational choice. Hence, Rawls explains that

> the principles regulative of the kingdom of ends are those that would be chosen in this position, and the description of this situation enables us to explain the sense in which acting from these principles

---

73Sandel, p. 120.

> expresses our nature as free and equal rational persons.74

Sandel argues, however, that Kantian metaphysics are implied because he believes that, on Rawls' account, the justification for the procedure must come prior to the two principles. In other words, justification for the description of the original position must be prior to its product.75 In support of his position Sandel emphasizes Rawlsian notions such as the idea that the original position reflects pure procedural justice. He also claims that as a procedural rendering of Kant's ethical view procedure must be prior. As a process ensuring pure procedural justice, the original position is a fair or correct procedure thereby guaranteeing that the principles that result are also correct or fair.76 He claims that for Rawls the justification of the procedure is prior to the principles because the principles

> ...may be conceived as principles that would be chosen by rational persons, and that in this way conceptions of justice may be explained and justified.77

Sandel maintains that the priority of procedure in fact confirms the Kantian self that he wishes to attribute to Rawls. He argues that

> the priority of procedure in Rawls' account of justification recalls the parallel priorities of the right over the good, and of the self over its ends. It thus connects the account of justification with the theory of the person which justice as fairness

---

74A Theory of Justice, p. 256.
75Sandel, p. 119.
76Sandel, p. 121.
77Sandel, p. 120

> was seen to entail, and suggests the importance of contract theory to the deontological project generally. As the self is prior to the ends it affirms, so the contract is prior to the principles it generates. Of course, not just any contract is prior to the principles of justice; as we have seen, actual contracts cannot justify precisely because they are typically situated in the practices and conventions which justice must assess. Similarly, real persons ordinarily conceived as 'thick with particular traits', are not strictly prior with respect to their ends, but are embedded in and conditioned by the values and interests and desires from among which the 'sovereign' self qua subject of possession, would take its purposes.78 (my emphasis)

Just as Sandel misinterprets the primacy of justice he is confused about the justification process involved in Rawls' theory. For Rawls the justification for the procedure does not wholly come before the two principles of justice. Justification for the conditions that are embedded in the procedure are in part based on a widely accepted and weak assumptions such as, the circumstances of justice, the moral arbitrariness of our assets, the assumption that the parties are rational and that they have personal conceptions of the good and life plans they wish to promote. These assumptions reflect an agreement that the principles should be chosen under certain conditions that ensure fairness. In other words, since the principles are the object of rational choice there are certain restrictions that must be followed if our reasoning is to produce determinate principles that are fair.79 But for Rawls the justification for our description of these conditions and restrictions also

---

78Sandel, pp. 120-121.
79A Theory of Justice, pp. 18-19.

depends on our considered convictions about justice. If the description of the original position does not match our judgements then we can either revise our judgements or the description of the original position.80 Sandel's error is in failing to recognize that although the principles are based on procedure for their derivation they must also express our considered convictions about justice. Hence, whether the principles cohere with those judgements or not will also provide part of the justification for the description of procedure.

Rawls states clearly that since the principles are not to be viewed as necessary truths, "a conception of justice cannot be deduced from self-evident premises or conditions on principles;...."81 Given that the principles are not to be viewed as necessary truths but rather the most reasonable ones that are chosen under certain conditions it seems natural that Rawls relies on a coherence view of justification. His argument cannot be a priori. If the principles are to accurately represent our sense of justice they must cohere with our considered convictions about justice. For Rawls, pluralism and justification based on consensus make it impossible for a general and comprehensive philosophical,moral or religious doctrine to "...assume a role of publicly acceptable basis of political justice."82

80A Theory of Justice, p. 20.
81Ibid, p. 21.
82John Rawls, "The Idea of an Overlapping Consensus," in Oxford Journal of Legal Studies, (Oxford: Oxford University Press), P. 6.

This, however, does not commit Rawls to the controversial claim that there is no independent moral order; rather he avoids this question altogether.83 Furthermore, pure procedural justice is not jeopardized because even though the principles must reflect our judgements of justice the original position as a procedure disallows irrelevant considerations such as knowledge of our assets, circumstances and full conceptions of the good. Hence, Sandel ignores the fact that justification for the original position is based on a coherence theory. Therefore, his argument for the Kantian ego interpretation is unsound because it must falsely assume that the procedure is strictly prior to the justification of the two principles.

In addition, the dispossessed view of the self does not follow from Rawls' theory because a procedural rendering of Kant eliminates metaphysical concerns. This is possible because as a procedural rendering the nature of the original position is to be viewed as strictly hypothetical.84

For example, Sandel claims that Rawls' political vision is flawed in part because Rawls is unable to avoid

---

83Ibid, p. 13.

84I wish to emphasize that this is not a general claim that hypothetical claims cannot entail metaphysical claims but rather that the way in which the original position is interpreted as hypothetical cannot admit metaphysics in the way Sandel suggests. This clarification is in response to an objection made by Steven DeHaven. The objection is that hypothetical claims can entail metaphysical claims. For example, if we make a hypothetical statement such as "If Jones were to ..." we will presumably be counting on Jones retaining certain traits. If the traits being retained are metaphysical then in this situation metaphysics is a relevant consideration.

metaphysical embarrassment. Sandel is convinced that the original position recreates "...the disembodied subject it resolves to avoid."[85] But if the original position is hypothetical is this true?

Rawls, I believe, is aware of the confusion that can arise from the original position. He warns the reader,

> ...not to be misled,..., by the somewhat unusual conditions which characterize the original position. The idea here is simply to make vivid to ourselves the restrictions that it seems reasonable to impose on arguments for principles of justice, and therefore on these principles themselves. Thus, it seems reasonable and generally acceptable that no one should be advantaged by natural fortune or social circumstances in the choice of principles. ...To represent the desired restriction one imagines a situation in which everyone is deprived of that sort of information.[86]

Furthermore, Rawls asserts that we can simulate the reasoning process by keeping before our minds the constraints that are represented by the original position, and providing arguments for principles in accordance with those constraints.[87]

Although Rawls does not require us to imagine ourselves as people in the original position devoid of any empirical attributes, we must be capable of simulating the reasoning process. We must be capable of identifying and rejecting arguments that involve designing the principles of justice on the basis of those of our attributes and characteristics that are arbitrary from a moral point of view. The original position and the beings that participate in it are purely

---

85Sandel, p. 14.
86A Theory of Justice, pp. 18-19
87Ibid.

hypothetical constructs and represent a reasoning process. In other words, the description of the original position is introduced as a heuristic device to clarify certain aspects of how we are to reason in order to arrive at principles of justice. Hence, even though we are required to simulate the reasoning process, we do so not by imagining ourselves as people in the original position (i.e., not by imagining ourselves without the particulars about ourselves and our situations), but rather by restricting the use of information about our particular situation and ourselves when we examine various arguments for and against principles of justice.88 Since the people in the original position are

---

88Given this account of the original position one might wonder whether Rawls could consistently maintain that there are some empirical attributes that are essential to personal identity. There are three views that could serve as candidates here. First Rawls could claim that some empirical properties are essential to personal identity but that knowledge of those essential properties is not itself essential to personal identity. Kripke, for example, suggests that one's origin may be essential to oneself, but also suggests one's identity is not affected by not knowing one's origins, whether through amnesia or postnatal kidnapping.(See Naming and Necessity, pp. 106-16.) This view would allow for one to engage in the reasoning process of the original position without creating problems concerning personal identity. Second, Rawls could claim that some cognitive accessibility to essential properties is essential for personal identity but that we can disregard that knowledge when we simulate the reasoning process. This view might require that Rawls think that some hypothetical reasoning is reasoning from inconsistent premises, and look to logics which lay out conditions under which this is legitimate. Moreover, Rawls would have to argue that the conclusions reached when ignoring those essential features of our identity can nonetheless affects us to act.. Third, Rawls could claim that the subject in my reasoning when I engage in hypothetical reasoning from the original position is not, strictly speaking me, but rather my 'counterpart' in that hypothetical situation. David Lewis suggests that the things in the actual world are bound to the actual world, but in counterfactual reasoning we reason concerning

theoretical constructions, and theoretical constructions do not entail metaphysical consequences, it is no objection to claim that they are not ordinarily conceived as 'thick with particular traits'. In other words, we are required to disregard our particular situations when determining the principles of justice. Yet, to do this does not mean that we must be people that lack those traits nor does it require that all my attributes are metaphysically contingent. For Rawls, readers of A Theory of Justice may be ordinarily conceived, for even though some of one's attributes may be metaphysically contingent this does not entail that all empirical properties are metaphysically inessential to one's personal identity. This is clearly Rawls' intended interpretation of this process for he states that when

---

'counterparts' which have features as close as possible to those in the actual world, but not all features. (Otherwise the possible worlds under consideration in the hypothetical reasoning would not be contrary to fact worlds.) The first and third alternatives seem to be the most promising options for Rawls because the second would require some argument in defense of the claim that even though some sense of identity is essential, this can be disregarded in hypothetical reasoning and more importantly the reasoning can still affect our actions.

However, Sandel himself must adopt one of these sorts of views in order to explain how it's possible for individuals (or communities) to engage in counter-factual reasoning. Reasoning about what would happen to me, my community and my world in various possible (but non-actual) situations is essential to deliberation and science. The decision to take out an insurance policy or to adopt certain social policies typically involves some such reasoning.

It's important here to distinguish the question of whether Rawls' hypothetical reasoning is morally relevant from the question of whether and how it makes sense. The first issue is what Taylor called the issue of advocacy in the communitarian/liberal debate; the second is what he called the metaphysical and ontological question. (See fn. 1 above)

"...we simulate being in this position, our reasoning no more commits us to a metaphysical doctrine about the nature of the self than our playing a game like Monopoly commits us to thinking that we are landlords engaged in a desperate rivalry, winner take all."89

An analogy to a jury may clarify the status of the people in the original position. The role of the veil of ignorance as a reasoning device is similar to what takes place when a judge disallows certain evidence that a jury has already been exposed to during a trail. The jury in this case is required to disregard this evidence in such a way that it is not used as grounds for deciding whether the accused is guilty or innocent.90 But to do this one does not have to attempt to erase this evidence from his memory or imagine that he was never exposed to the evidence that must be ignored.

89"Justice as Fairness: Political not Metaphysical", p. 239.
90 One may object to Rawls account on the grounds that he assumes that our desires and hence action must be the slaves of reason, insofar as the reasoning in the original position must determine how we act. It may also be argued that in this sense Rawls' theory is Kantian because individuals' desires and actions must be capable of being determined by reason alone. However, this objection is unfounded because it neglects the role of reflective equilibrium and Rawls' coherence theory of justification. Although Rawls must claim that our desires and especially our sense of justice, and thus our actions, may be affected by engaging in the reasoning process from the original position, he also thinks that our construal of the original position and our deeming it relevant to questions of justice are influenced by our considered convictions about justice. Hence, our actions and desires are not solely grounded in Kantian 'pure' reason.
91 I owe this analogy to Steven DeHaven.

I show thus far, that Sandel misinterprets Rawls theory in at least two ways. First, Rawls need not claim that the principles of justice must be derived independently of all societal values in order for his theory to be deontological. Second, Sandel claims that the justification of procedure in Rawls' theory must come prior to the principles of justice. But, this is not so for Rawls because he relies on a coherence theory of justification which requires that the justification of procedure is in part dependent on whether it coheres with our considered convictions of justice. Finally, even though I do not give a conclusive proof that the Kantian ego is not entailed by Rawls' theory, we should be suspicious of Sandel's claim insofar as two of the arguments this view of the Rawlsian self is derived from involve a serious misreading of Rawls. So I show at least that these two arguments fail to demonstrate that the Kantian ego follows from A Theory of Justice.

CHAPTER II

Section 1. Logical, Conceptual and Epistemological Priority

Sandel's main concern is to expose what he believes is the Rawlsian theory of the person. It is a theory in which as selves "..., we must stand to our circumstances always at a distance, conditioned to be sure, but part of us always antecedent to any conditions."1 In this chapter I shall distinguish between what I call logical priority, conceptual priority, and epistemological priority. Then I shall argue that if the priority Sandel attributes to the Rawlsian subject is any of the three I present then this priority is not damaging to Rawls' theory. I shall also establish that the Kantian ego is prior to its ends in that all the empirical features of the subject are contingent to personal identity.2 By identifying the Rawlsian self with the Kantian ego Sandel attributes the same interpretation to Rawls. However, this conception of the self does not follow from Rawls' view.

Although Sandel claims that the proposition "the subject is identified prior to its ends" reflects some thesis of epistemological priority he also suggests that the subject is in some sense logically prior to its ends.

1Sandel, p. 10.
2Presumably this subject requires some features in order to be viewed in space and time.

Sandel asserts that the "...description of this subject will have a distinctive logical status. It will in some sense be necessary, non-contingent, and prior to any particular experience-...".1(my emphasis) But what is the distinction between epistemological and logical priority?

I take logical priority to be a relationship between propositions. Where P and Q are propositions, P is logically prior to Q only if Q entails P but P does not entail Q. The priority is represented by the one way entailment between P and Q. It may be the case that Sandel interprets the priority and independence of the subject to its empirical features in this logical sense. For instance, to say that Graham is logically prior to his empirical property of wearing a red hat is to claim that Graham's wearing a red hat entails that Graham exists. But asserting that Graham exists does not entail that Graham necessarily exists wearing a red hat, just because Graham's wearing a red hat may be a contingent property of him. Furthermore, the logical priority of Graham's wearing a red hat does not entail that Graham exists without any empirical properties. It may entail that some empirical properties are contingent but it does not entail that all empirical properties are contingent. (It's worth noting that even the latter does not entail

---

1Sandel, p. 49.

that Graham could exist without any empirical properties at all, ie. as a disembodied subject.)

Conceptual priority is somewhat different. Where P and Q are concepts, P is prior to (P and Q) if and only if (P and Q) includes P but P does not include (P and Q). Conceiving of Graham wearing a red hat is to conceive of Graham, but not vice versa. Thus Graham is conceptually prior to his wearing a red hat. This conceptual priority may entail that some empirical properties are contingent but it does not follow that all such properties are contingent.

Epistemology in general is concerned with providing conditions for justifying beliefs to be true for knowledge. Epistemological priority can be defined as follows: Where P and Q are propositions, P is epistemologically prior to Q if and only if one cannot have a justified belief in Q without also being warranted in believing P to be true, but one can be warranted in believing P independently of believing Q. If I am justified in believing that Graham is wearing a red hat, I am justified in believing that Graham exists, but not vice versa. The epistemological priority of Graham's existence to his wearing a red hat does not entail that all empirical properties are contingent to Graham.

If the priority of the self is either logical, epistemological or conceptual priority, then the metaphysical consequences for Rawls are insignificant.

It is doubtful that Rawls disagrees with any of the true claims that follow from these interpretations but these claims are not metaphysically problematic for Rawls. However, I show in the next section that Sandel believes that the conception of the self that he believes follows from Rawls' view is one in which all features of the subject are contingent to personal identity. It is this claim that I also show need not follow from Rawls' theory. My objection is in part substantiated on the grounds that in order for Sandel to object to Rawls' theory on the basis of dubious metaphysics Sandel must interpret the Rawlsian self as having significant metaphysical implications. Hence, it is unlikely that Sandel interprets the self's priority as prior in any of the three senses I mention because all the metaphysical implications of understanding its priority in these ways are banal.

Section 2. The Kantian Self

It is important to examine Kant's view of the self and Sandel's interpretation of it because Sandel identifies the Rawlsian subject with the Kantian self. Here, I show that Sandel's interpretation of the Kantian self is specious. Hence, even on Sandel's own grounds Rawls is not committed to the view that Sandel attributes to Kant because it is based on a mistaken interpretation of Kant. Moreover, even if a plausible view of the Kantian self cannot be given - a question I do not propose to explore here, Rawls is not committed to this conception because he maintains that

metaphysics is independent of normative theory. Hence, all that Rawls is required to do is to adopt a plausible view of the self that is compatible with his normative theory.

There are three features of the Kantian self. The first feature is the transcendental unity of apperception. It is expressed by the phrase "I think". It is the form of consciousness or thought which is comparable to the Cartesian "Cogito". In isolation, the "I think" is the formal and logical structure of thought which can only be viewed as subject in a subject\predicate relation. The unity of apperception cannot by itself be experienced because it is subject and cannot be taken as an object (i.e., it lacks empirical content). However, one can recognize that one is aware or conscious in virtue of the application of a predicate. As Kant explains, the "I think",

> ...in itself (is) completely empty,...; and we cannot even say that this is a concept, but only that it is a bare consciousness which accompanies all concepts. Through this I or he or it (the thing) which thinks, nothing further is represented than a transcendental subject of the thoughts=X. It is known only through the thoughts which are its predicates, and of it apart from them, we cannot have any concept whatsoever,....1

If the transcendental unity of apperception is not self-knowledge what is such knowledge and how do we acquire it? The answer to this question will reveal the other two features of the Kantian self.

1Immanuel Kant, Critique of Pure Reason, ed., Norman Kemp Smith, (London: Macmillan), p. 331.

Even though the "I think" has no empirical content it is the vehicle through which we acquire self-knowledge. The transcendental unity of apperception synthesizes and unifies our experience. Hence, self-knowledge is knowledge of those of our empirical properties that are acquired through introspection and sensory awareness. This is the second aspect of the self which is termed phenomena. Phenomena are properties such as our beliefs, values, and how we appear through sensory awareness (i.e., phenomenon is the self as it appears).2 In virtue of the unifying and synthesizing function of the "I think" it is possible to have knowledge of our empirical properties because the subject takes our experience as an object. For example, the subject is nothing but consciousness but when I predicate it by adding the phrase "I have brown hair" I am representing "...an intuition of the manifold in me, by which I determine my thought."3 In short, the transcendental unity of apperception is the potential for the unification of our experience via determinant judgement.4

The noumenal feature of the self is the "thing in itself" (i.e., the essential properties of the self) that cannot be known empirically but nonetheless must be assumed.5 Kant believes our empirical properties must be grounded in essential properties because the former properties are nothing by themselves outside of our

2Kant, pp., 168-169.
3Ibid.
4Kant, p. 329.
5Kant, pp., 267-268.

representation of those properties.6 The noumenon can only be employed in negative terms because its constitution cannot be known. Hence, what we mean by noumenon is that it is "...not an object of our sensible intuition,...".7 Kant explains a second reason why we must assume that there are essential properties underlying our empirical properties when he states that, man

> ...knows himself also through pure apperception; and this, indeed, in acts and inner determinations which he cannot regard as impressions of the senses. He is thus to himself, on the one hand phenomenon, and on the other hand, in respect of certain faculties the action of which cannot be ascribed to the receptivity of the sensibility, a purely intelligible object.8

For Kant the representation of our consciousness is a thought and not an intuition, hence it is not attributable to our sensory awareness. Since we actively synthesize and unify our experience we must think of our selves as noumena. The noumena and phenomena are similar in that they are to be considered as objects whereas the transcendental unity of apperception can only be considered as subject. If the subject of apperception

> cannot be regarded as an object in the world, then it cannot be equated with the noumenal self, for the concept of the latter is the concept of an object in the world (the "transcendental-object of inner sense"), albeit one that can be known only in a unique nonsensible manner.9

6Kant, pp., 269-270.
7Kant, p. 268.
8Kant, p. 472.
9Henry Allison, Kant's Transcendental Idealism: An Interpretation and Defense, (London: Yale University Press), p. 291.

Kant's motivation for making the noumena-phenomena distinction is in part a response to the problem of how to resolve the incompatibility of freedom and natural necessity. Kant believes that by viewing a person in two different ways (i.e., as phenomena and noumena) we can account for our freedom when we view ourselves as moral agents and if we view ourselves as empirical we can account for the fact that we are governed by the laws of nature.10 Freedom as an intelligible cause is a transcendental idea, and hence it cannot be employed empirically. Therefore Kant must posit a notion of freedom in a positive practical sense which is freedom as self-legislation. It is empirical man that must be posited as free. Kant clearly states that "...the concept of freedom is the final purpose which (or its phenomenon in the world of sense) ought to exist, and the condition of the possibility of this is presupposed in nature (in the nature of the subject as a sensible being, that is, man)."11(my emphasis)

Sandel's characterization of Kant's view is somewhat different from the one suggested. First Sandel identifies the transcendental unity of apperception as the subject that unifies and synthesizes our experience and that is "...antecedent to any particular experience...."12 He then claims that we can view ourselves as both an object and a subject of experience which leads to a practical argument

---

10Kant, p. 472.
11Immanuel Kant, Critique of Judgement, translated by J.H. Bernard, (New York: Hafner Publishing Company), p. 33.
12Sandel, p. 8.

which justifies the priority of the subject over its empirical properties.

> Qua object of experience, I belong to the sensible world; my actions are determined, by the laws of nature and the regularities of cause and effect. Qua subject of experience, by contrast, I inhabit an intelligible or super-sensible world; here, being independent of the laws of nature, I am capable of autonomy, capable of acting according to a law I give myself.13

From my previous discussion we can see that the intelligible feature of ourselves is the noumenon which for Kant is not the subject of experience as Sandel suggests but rather the transcendental object. Only the transcendental unity of apperception is to be taken as the subject of experience which is not to be identified with the noumenal self. Hence it appears that Sandel confuses the two, for in the previous quotation he identifies the intelligible object (i.e., the noumenon) with the subject of experience (i.e., the transcendental unity of apperception). By so doing Sandel suggests that the noumenon is prior to the phenomenon. Yet, priority is an attribute of the transcendental unity of apperception and not the noumenon.

These claims made by Sandel lead to the first of two possible interpretations that he might have in mind in his discussion of the Kantian ego. We must remind ourselves that the Kantian ego must be considered from a metaphysical point of view as opposed to a moral point of view if we are to find a <u>metaphysically</u> problematic account. By confusing the transcendental unity of apperception with the noumenon

---

13Sandel, p. 9

Sandel believes that the noumenon is prior to the phenomenon. However, if we consider the sense in which the "I think" is prior to its predicates I show that on this interpretation the metaphysical consequences are banal. Apperception is "bare consciousness" and should be viewed as the activity of thinking rather than a thinker. The process of thought or conscious activity is logically prior to the contents of the mind (i.e., concepts or thoughts) insofar as when I think of Graham as having blond hair this entails that thinking is going on or that I am conscious. But the priority of the activity of thinking does not entail that thinking exists or is experienced independently of the concepts or predicates that are its objects. In fact Kant, as we see, explicitly states that "It is known only through the thoughts which are its predicates, and of it apart from them, we cannot have any concept whatsoever,...."[14] Nor does it entail that, to be engaged in the activity of thinking that I must be thinking of Graham with blond hair; what is entailed, rather, is that if I have a thought I must be thinking of something. Viewed in this way the priority of apperception does not seem to be metaphysically problematic because Rawls need not reject the true claims that follow. Hence, Sandel must have a different interpretation in mind.

Sandel, in the end equates the Kantian ego with the Rawlsian self. Hence, we may gain some insight into the way

---

14Critique of Pure Reason, p. 331.

the transcendental unity of apperception and the noumenon are confused by examining the metaphysically problematic view (I am assuming that Sandel must attribute this view to Kant) that he believes Rawls must avoid. Sandel believes that Rawls must avoid a self that is "...totally detached from its empirically given features..." because this "...would seem no more than a kind of abstract consciousness (conscious of what?), a radically situated subject given (sic) way to a radically disembodied one."15 (my emphasis) From this it is clear that Sandel believes that the noumenon is consciousness or the transcendental unity of apperception, as opposed to merely applying a notion of priority to the noumenon.16 Since Kant denies that we can know anything about the noumenon how we can know the noumenon as consciousness? More importantly, Sandel is committed to interpreting the Kantian self as one in which all empirical attributes are metaphysically contingent for

---

15Sandel, p. 21.

16Perhaps Sandel merely intends to attribute the notion of priority to the noumenon as the "thing in itself" rather than as consciousness. I am not sure what this account amounts to but I suspect it might be something like the following: If the noumenon (i.e., the essential properties of the self) are prior to the phenomenon (i.e., the empirical properties of the self) then to say that Graham is blond entails that Graham has essential properties that ground his presumably contingent empirical properties. But from this it does not follow that we must view Graham as having blond hair (perhaps he dyed his hair) or that we can conceive of Graham without any empirical properties. Even if we allow Sandel to apply a notion of priority to the noumenon it is unlikely that Rawls rejects the true claims that follow. More importantly the metaphysical consequences in this case are also innocuous.

the self or the subject as consciousness.17 There is additional evidence that suggests that this is Sandel's interpretation. For instance, Sandel and Rawls seem to share the view that

> the Kantian conception suffers from obscurity and arbitrariness, for it is unclear how an abstract, disembodied subject could without arbitrariness produce determinate principles of justice, or how in any case the legislation of such a subject would apply to actual human beings in the phenomenal world.18

Further Sandel claims that where Rawls attempts to divorce himself from Kant is "...in denying that a prior and independent self can only be a transcendental, or noumenal subject, lacking altogether an empirical foundation."19 Sandel believes that the metaphysically suspect nature of Kant's view is the following: if the noumenon is the relevant ethical agent how is it possible for freedom to be actualized in experience or the empirical world without arbitrariness? In short, if all empirical properties are metaphysically contingent to the noumenon how could the noumenon make determinate moral judgements in the empirical world? However, this need not be a problem for Kant because as I show, Kant's intention is to posit freedom in a positive sense as self-legislation that applies to man as a sensible being. If Kant is successful, then the moral agent

---

17This previous citation also suggests Sandel thinks noumena can exist independently of all empirical attributes. If Kant holds this view it is metaphysically problematic, but there does not seem to be sufficient evidence to firmly establish this much stronger claim as Sandel's intention.
18Sandel, p. 13.
19Ibid.

is capable of freedom and is empirically grounded contrary to what the latter citation suggests.

Sandel presents what he believes is the metaphysically problematic view of the Kantian self by claiming that for Kant the noumenal self is pure consciousness (i.e., it is the transcendental unity of apperception). Support for this second interpretation is suggested when Sandel states that

> the independence of the subject does not mean that I can as a psychological matter, summon at any moment the detachment required to overcome my prejudices or step outside my convictions, but rather that my values and ends do not define my identity, that I must regard myself as the bearer of a self distinct from my values and ends, whatever they may be.20

If our empirical properties cannot define our identities then presumably they are inessential for personal identity. If they are inessential they are contingent and could have been otherwise.

The most convincing evidence that this is the Sandelian interpretation of Kant comes when he discusses a problematic feature of Kant from a metaphysical point of view.

> ...Kant's doctrine requiring abstraction from <u>all contingency</u> may be unable to distinguish between the lives of the saint and the scoundrel, as long as both are lived by a consistent set of principles freely chosen and conscientiously acted upon. The choice of the <u>noumenal</u> self may-in fact might necessarily-be arbitrary in this sense.21 (my emphasis)

Sandel claims that the noumenon makes moral judgements independently from <u>all</u> contingency. Furthermore, Sandel believes that, for Kant, what is contingent is empirical.

20Ibid., p. 12.
21Sandel, pp., 37-38.

Hence, Sandel believes that for Kant all empirical attributes are contingent to the noumenon. In addition, Sandel claims that Kant's theory requires that we must abstract from all of our contingent empirical properties which makes us incapable of distinguishing the lives of the saint and the scoundrel. What is implied is that if we have access to our empirical properties we can individuate the saint and the scoundrel in virtue of empirical attributes of their respective lives. But absent consideration of those attributes the noumenon is all that remains and it is an inadequate criterion for individuation. If this is the case then we begin to see why Sandel believes that in Kant it "might necessarily" be impossible to distinguish the lives of the saint from the scoundrel. If this is true for Kant we are unable to distinguish between the scoundrel and the saint because we do not know the noumenon empirically and as human beings we do not have the capacity to intuit it in a nonsensible manner.22

22There is a sense in which Kant may not be able to distinguish the saint from the scoundrel. But this is so strictly from a moral point of view and not a metaphysical point of view. If the saint acts on a maxim because he is benevolent by nature and the scoundrel acts on a maxim because of his natural self-love then neither of these acts are morally worthy. And neither of these acts is done out of duty. However, since Kant posits empirical man as being a self-legislator it is not the case that when we apply the categorical imperative that we must choose as pure consciousness independently of our empirical properties. The reason Kant gives for rejecting a maxim is that it does not conform to the standards of rationality. The examples that he gives in "The Grounding for the Metaphysics of Morals" are rejected as maxims because they cannot be willed as categorical imperatives without also being self-contradictory. He does not reject maxims because the maxim includes reference to empirical ends and situations. Given

Sandel concedes that Rawls intends to bypass Kantian metaphysics but he also believes that Rawls is unsuccessful because his theory "...recreates in the original position the disembodied (Kantian ego) it resolves to avoid."23 If the Kantian ego is to represent an implausible metaphysical stance and force Rawls into metaphysical embarrassment I suspect that the second interpretation or one very similar must be Sandel's view of the Kant. In virtue of the fact that Sandel identifies the Rawlsian self with the Kantian self, then this interpretation must also apply to Rawls. Now the question that must be answered is whether this notion of the self in fact follows from Rawls' theory. It must be admitted that Rawls does says that the "...self is prior to the ends which are affirmed by it;..."24 But the sense of priority here is reflected by the claim: "...;even a dominant end must be chosen among numerous possibilities."25 Does this view of priority imply a view in which all empirical properties are contingent to the subject when taken from a metaphysical stance? I think not.

---

this, it does not necessarily follow from Kant's claims that we must apply the categorical imperative as selves that lack empirical properties (i.e., as noumenal selves). All that follows from this is that our maxim (which may include refernce to empirical properties i.e., desires, motives and ends) cannot be willed as a universal law then the maxim must be rejected because it conflicts with our duty.

23Sandel, p. 14.

24A Theory of Justice, p. 560.

25Ibid.

Section 3. The Priority of the Rawlsian Subject

For Rawls, the priority of the self or the subject to its ends is expressed by the claim that the subject is prior and independent of its ends because its ends are chosen as opposed to merely given. From a moral point of view the autonomy of the individual is reflected when the individual actively chooses his ends, his conceptions of the good and rational plans of life. The subject must be prior in this sense otherwise these ends cannot be chosen. From a metaphysical point of view, logical, conceptual, and epistemological priority may be applicable to his theory. For instance, if the subject is logically prior to the ends he chooses and the subject has ends the subject exists. But, in none of the three senses of priority does it follow that all empirical properties are contingent, or that the subject exists without any empirical properties. What Rawlsian priority means is no more than that one's choice to become a doctor, for example, may be an accidental feature of that individual which could have been otherwise. It is at this point unnecessary to examine the other senses of priority because they are also metaphysically innocuous.

There are two possible Sandelian interpretations of the Kantian self. On the logical priority interpretation the metaphysical consequences are unproblematic for Rawls; hence on these grounds this interpretation is rejected. The pure consciousness interpretation, suggests that all empirical properties are contingent to the subject and hence fulfills

the requirement of being metaphysically dubious. By identifying the Rawlsian self as the Kantian self Sandel reveals the view of the self that he believes is the basis for A Theory of Justice. However, the problem Sandel faces is that as described this view need not follow from Rawls' claims about the priority of the subject. Furthermore, if Sandel's strategy is correct then the original position may entail some theory of the subject. But, this theory must be a theory of a subject from a moral point of view. Furthermore, I argue that this view need not entail the metaphysics that are suggested by Sandel. Sandel here, derives a subject based on a metaphysical stance from a subject based on a normative stance which commits him to the mistake of conflating ethics with metaphysics.

Section 4. Individuation and Personal Identity

The previous fallacy may be symptomatic of a more serious error. Personal identity has been concerned with searching for an interesting relation such that a person x at time T and person y at T1 can be identified as the same person if they share that relation.26 Candidates for this relation have included such things as bodily continuity, continuity of memory and psychological continuity. It may even be the case that our search is in vain and no such criteria exist.27 When we speak of individuation we are

---

26Where R is a relation this may be formally expressed by the following: (N y=x iff xRy).

27Parfit makes claims similar to these in Reasons and Persons. I wish merely to point out that I am not proposing any particular view of personal identity.

concerned with features or relations that we use to decide whether x is or is not the same as some y. It might be the case that a relation can be used to show that x is in fact the same as y and hence can be used to determine identity. But to use a feature or relation to individuate in this sense does not in and of itself require that it need be one involved in establishing an identity relation.

If distinctions are not made between the three senses of priority and the notion of priority that Sandel attributes to Rawls, the issues of personal identity and individuation become confused. More specifically, conflating this distinction results in asserting that empirical properties that are used merely to distinguish between individuals must also be viewed as contingent or inessential properties for personal identity. For instance, in the case of logical, conceptual, and epistemological priority saying that the subject is prior to its ends does not imply any preferred metaphysical conception of the self. These concepts are compatible with both the view that no empirical properties are essential, and the view that at least some empirical properties of the self are essential. The salient feature of these concepts is that on their own they do not provide a conclusive metaphysical view of the self.

However, they may provide a means for distinguishing individuals from one another. For instance, in all three cases of priority, one can individuate subjects from each

other on the basis of their attributes and in doing so, nothing need be implied about the metaphysical status of those attributes in connection to the identity relation. In other words, distinguishing individuals on the basis of attributes does not require that these attributes must be contingently related to personal identity or that alternatively, they must be essential to personal identity. On the other hand, Sandel's interpretation of Rawlsian priority does have important metaphysical significance for identity. Sandel believes that saying that there must be a self that is identifiable prior to and independently of its empirical properties is making a claim about the nature of persons and personal identity. This sense of priority reflects the negative view that the identity of the self is not determined by its empirical properties as opposed to providing a positive view which would list the necessary and sufficient conditions for the identity of the self. Hence, to make the claim that there is a self independently of its ends is to articulate the claim that these ends do not serve as necessary conditions for personal identity. In other words, these criteria of individuation are metaphysically contingent. Hence, Sandel believes that the view that all empirical properties are contingent is entailed by a Rawlsian sense of priority. If he affirms this claim then the attributes of the subject that may be used to individuate one individual from another must be viewed by him as inessential properties for personal identity.

Conversely, if the entailment is not affirmed the metaphysical status of properties for individuation in relation to identity are indeterminate. Some may turn out to be either essential or inessential properties. Herein lies Sandel's mistake.

In the former case the entailment precludes the possibility that there may be some empirical properties of the subject that are essential for personal identity. There is no reason to suppose that this is in fact the case. By looking at his argument and a formal reconstruction the fallacy will be exposed. Sandel informs us that one essential element of Rawls' theory is the notion of plurality. In other words, for justice to be primary there must exist more than one person. In addition, he claims that for subjects to be plural there must be some way of individuating persons that is, there must be some way of "...distinguishing one from another,...."28 Sandel also suggests that this is done via empirical properties because no two people are in every respect identical.29 However, he then claims

> as the account of plurality suggests, not just any subject of possession will do, but only an antecedently individuated subject, the bounds of whose self are fixed prior to experience. To be a deontological self, I must be a subject whose identity is given independently of the things I have, independently, that is of my interests and ends and my relations with others.30

28Sandel, p. 51.
29Ibid.
30Sandel, p. 55.

I show that if the Kantian ego, as it is interpreted by Sandel, is to follow from Sandel's interpretation of Rawls' view of plurality an implicit false premise must be present in his argument. Moreover, Sandel's argument is based on confusing personal identity and individuation.

1. Rawls' theory requires a plurality of individuals.

2. These individuals are distinguished by a principle of individuation.

3. The principle of individuation for Rawls, is based on empirical features such as mental and physical characteristics.

4. However, these properties need not be essential to personal identity and in fact they are often accidental or contingent to the individuals that we are.

5. If we are individuated by empirical properties some of which are accidental it follows that all empirical properties are contingent to our identities.

Therefore, this account of plurality suggests that a deontological ethic requires a self in which all empirical attributes are contingent to personal identity.

When we wish to distinguish individuals we use empirical properties that are not common to those that are being distinguished. But the nature of individuation does not require that these properties are necessary for personal identity. Hence, premise 4 reflects the indeterminate metaphysical status of these properties. The fifth premise is implicit in Sandel's argument because only on the basis of 5 does the conclusion follow. The problem for Sandel is that the fifth premise that must be assumed is also clearly false. It does not follow that, if some empirical

conditions that are employed for individuation are accidental, all empirical conditions are inessential to personal identity. Sandel's argument falsely precludes the possibility that any empirical conditions can constitute necessary conditions for personal identity by conflating the problem of individuation and the problem of personal identity. The notion of plurality allows that some empirical properties may be accidental but it does not entail or suggest that all empirical criteria are accidental or contingent to personal identity. Once the distinction between the two problems is made, it is clear that deontological ethics does not require a self that exists independently of its ends.

Section 5. The Epistemology and the Metaphysics of the Self

Sandel objects to Rawlsian reflection or self-knowledge on the grounds that it cannot "...take as its object the self qua subject of desires....It takes as its objects the contingent wants and desires and preferences of the self, but not the self itself."[31] With further explanation I show that this objection represents Sandel's confusion between issues involving the epistemology of the self and issues concerning the metaphysics of the self. This is similar in kind to his conflation of personal identity and individuation.

Sandel's objection to Rawls' view of self-knowledge only serves to exemplify his confusion. Sandel contends

---

31Sandel, p. 159.

that Rawlsian self-knowledge is a consideration of a person's desires, interests and preferences. Reflection on Sandel's reading is merely an awareness of our accidental properties. Sandel states that

> ...even if 'our direct self-knowledge' permits some uncertainty for reflection to sort out, the self that is known once the uncertainty is resolved is not really the self in the strict sense distinguished throughout but merely the contingent accidents and attributes of the self.32

This view may represent at least one feature of self-knowledge that Rawls can consistently adopt. For instance, for Rawls, the epistemology of the self may be a consideration of how we identify and re-identify the self as the same self over time. In other words, an epistemological question concerning the self might involve a question like: How do I know (i.e., how can I be justified in believing) that I am the same self now as the self that had breakfast at 8 a.m. this morning? The answer to this question may depend on introspection, perception, and memory. Sandel's objection suggests that self-knowledge on Rawls' view may be in part arrived at through introspection. This is a feature of how we acquire justified true beliefs about the identification of our selves as the same selves over time. But this account has less philosophical significance for Rawls' theory than Sandel suggests. By recognizing my accidental properties I have at a given time, I may use them as a way of picking out my self as that same self at a later time. I in effect distinguish myself from others on the

32Sandel, p. 160.

basis of mutually exclusive accidental properties. This is similar to distinguishing different individuals from others at a particular time via different accidental properties.

If this is Rawls' view of self-knowledge then it is merely a consideration of our accidental properties that may be used as a factor in distinguishing ourselves from others. He is not concerned at this point, about questions of personal identity. Hence, to even initiate the objection on Sandel's part is to change the topic from epistemological questions of the self to metaphysical questions about personal identity. But if Sandel's complaint is that self-knowledge should be concerned with the "self" qua subject of desires rather than the desires of the self (i.e., the accidental properties of the self) then we should be concerned with the essential properties of the self. Herein lies Sandel's confusion. Issues involving the self and epistemology deal with questions about what sorts of things I can know about myself in order to recognize myself as the same self through time. Metaphysical questions about personal identity are concerned with identifying the self as the same self by determining the essential properties of the self. Here, Sandel equivocates between self-knowledge qua justified true beliefs and self-knowledge qua knowledge of essential properties of the self. Hence, Sandel's objection is inappropriate to the topic of discussion.

We may agree with Sandel that Rawls' discussion of self-knowledge cannot be a consideration of the essential

criteria for determining personal identity. But, we do so not on the grounds that "...Rawls' self is conceived barren of constituent traits, possessed of contingent attributes held always at a distance,..."33 as Sandel suggests. Rather our agreement with Sandel is explained by showing that Rawls' consideration of self-knowledge is a different issue from a consideration of essential properties of personal identity. In other words, to explain what Sandel perceives as a limitation in Rawls' view, as representative of the Kantian ego, is to misread Rawls.34

One might wonder why the distinction between epistemological and metaphysical issues concerning the self should be made. If we denied this distinction we would be incapable of accounting for trivial cases of re-identifying the self over time without also knowing the essential properties of the self. As an illustration, we can borrow a simple example from Kripke. If we are trying to identify a table in ordinary circumstances we do so by properties such as its woodenness, brownness and that it is in the room. But by doing so we may be discussing its accidental properties and not its essential properties. As Kripke suggests, this should not worry us because

> some properties of an object may be essential to it in that it could not have failed to have them.

33Sandel, pp. 160-161.

34It must be emphasized that even though Rawls does not consider the nature of personal identity, this does not preclude the possibility that he can develop a view of personal identity that is compatible with his political view, and yet different from the view that Sandel attributes to him.

> But these properties are not used to identify the object in another possible world, for such an identification is not needed. Nor need the essential properties of an object be properties used to identify it in the actual world by means of properties.35

It is implausible to view every case of identification to depend on identifying essential properties. This is certainly one reason why we should distinguish questions about epistemology and personal identity.

In sum, I have shown that Sandel misinterprets Rawls' view that justice and procedure are primary. He also misinterprets the way in which the subject must be prior to its ends for Rawls. If Sandel identifies the priority of the Rawlsian subject with either logical, conceptual or epistemological priority, then Rawls' view is not metaphysically problematic. On the other hand, if we accept the interpretation that is metaphysically problematic it simply does not follow from A Theory of Justice. And by so doing he also confuses individuation with personal identity and he confuses questions of epistemology with metaphysics. In the following chapter I argue that the conceptions of the self that are derived from Rawls' ethical claims about the moral arbitrariness of our assets and that we ought to treat our assets as common assets are fallacious. Sandel's error is in deriving metaphysical conclusions from ethical premises. Sandel is also faced with the dilemma of either

---

35Kripke, Saul. Naming and Necessity. Basil Blackwell, Oxford, 1980, pp, 52-53.

viewing the Rawlsian self from a metaphysical or normative point of view.

CHAPTER III

Here I wish to discuss two separate but related issues. First I show how claims for both the Kantian conception of the self and the constitutive conception of the self (i.e., the wider notion of the self) are founded upon fallacious arguments. I contend that Sandel erroneously derives metaphysical conclusions from ethical premises. The transcendental ego is derived from Rawls' claim that our assets and the benefits that flow from them are not deserved which is the third argument for this conception of the self.1 Hence Rawls maintains that they should not play a role in our reasoning for the principles of justice. The constitutive conception of the self is derived from Rawls' statement that if our assets are morally arbitrary then we ought to treat our assets as if they are common assets. On the surface, Sandel's arguments seem sound. However, once the status of these premises and the conclusions are compared his error becomes apparent. For Rawls both these claims are moral premises, yet the conclusions reached by Sandel are metaphysical. As moral premises they imply no preferred metaphysical conception of the self. Thus, although the Kantian self is not incompatible with Rawlsian moral claims it also does not necessarily follow from his

---

1The other two arguments are based on the primacy of justice and the primacy of procedure that are discussed in chapter I.

theory. The constitutive conception of the self, however, is diametrically opposed to the individualistic project of A Theory of Justice. The metaphysical conceptions of the self that Sandel imputes to Rawls are undermined because Sandel is unable to affirm an entailment between the conclusions of his derivation and the premises from whence they came. If the entailment cannot be affirmed then there is no compelling reason to suppose that these conclusions are the preferred metaphysical conceptions of the self. Second, I argue that the means-ends dilemma that Sandel and Nozick pose for Rawls can be resolved. In brief, Sandel contends that Rawls is confronted with three choices if he wishes to retain the difference principle as a distributive principle. First, Rawls may accept the difference principle as it stands but then Sandel argues in defense of Nozick that he is committed to treating people as means as opposed to ends. Sandel maintains that since the difference principle embraces the claim that we should treat assets as common assets then this is tantamount to using people as means by using their assets. Sandel also believes that without an appropriate metaphysical foundation this objection cannot be avoided. Therefore, Sandel further denies that this is a viable alternative for Rawls. The second alternative that is posed for Rawls is to adopt the Kantian disembodied self as the metaphysical foundation for the difference

principle. This conception of the self rescues Rawls from the means-ends objection by separating the self from its ends in such a way that one's assets are viewed as inessential to personal identity. If our assets are inessential then they cannot be viewed as assets qua self and hence using them as common assets does not violate the respect for persons. But Sandel adds that this comes at the cost of accepting a metaphysically suspect view of the self and thus, the second alternative is equally problematic. The final alternative is to accept the constitutive conception of the self. As a metaphysical foundation the self properly applies to the community as opposed to the empirically individuated human being. Hence, individuals cannot object that they are being treated as means as opposed to ends in themselves because the concept self qua individual does not apply. To treat assets as if they were common assets is for Sandel treating selves as ends because the self is now the extended notion of the community. Sandel correctly informs us that this alternative is in general the least attractive to Rawls because accepting this view is contrary to his individualistic project.

Section 1. The Kantian Ego Fallaciously Derived

In chapter I I rejected Sandel's argument that the transcendental ego follows from the priority of procedure and justice. Here, I shall show why the remaining argument is fallacious. Sandel derives the disembodied

conception of the self, in which all its empirical properties are contingent, from the Rawlsian premise that our assets are undeserved and hence they are morally arbitrary. For Rawls, the view that our wants and desires are contingent is based on his description of the original position and the justification for the veil of ignorance. The argument goes something like the following. When one is choosing principles of justice that are to serve as principles for evaluating institutions they must be chosen independently of our particular conceptions of the good but of course not independently of our considered convictions about justice. Moreover, our natural assets and social positions are contingent, a matter of accident, and they are "morally" arbitrary because we did nothing to deserve them. To base the principles of justice on facts that are morally arbitrary allows individuals to tailor the principles of justice to their advantage. This in turn, nullifies the requirement of deriving the principles of justice independently of people's particular desires and interests.[1] In short, justice must be prior to conceptions of the good.

Since our attributes and social positions are arbitrary from a moral point of view our particular conceptions of the good should play no role in the choice of the principles of justice. In other words, our assets and social positions are irrelevant, and must be excluded as reasons for accepting one principle of justice over another. Sandel

---

[1] A Theory of Justice, p. 12.

mistakenly believes that if our ends are arbitrary from a moral point of view then they are also contingent from a metaphysical point of view. Sandel claims that deontological liberalism requires us to see ourselves as Kantian egos. Sandel's error is to think that the fact that empirical properties should not be taken into consideration from a moral point of view entails that empirical attributes are metaphysically contingent. I suggest that this move is attractive to Sandel in part because the people in the original position are to disregard their particular circumstances. Hence Sandel views these people as transcendental selves. But for Rawls, to say that one's attributes are morally arbitrary is just to deny them any justificatory force when arguing for principles of justice. It does not mean, contrary to Sandel's opinion, that these attributes are metaphysically contingent. However, before this can be established an examination of Sandel's specific argument is critical.

Sandel believes that Rawls is committed to this theory of the person if he also wishes to undermine desert. He claims that

> We can see in this light how Rawls' argument from moral arbitrariness undermines desert not directly, by claiming I cannot deserve what is arbitrarily given, but indirectly, by showing I cannot possess what is arbitrarily given, that is, 'I' qua subject of possession, cannot possess it in the undistanced, constitutive sense necessary to provide a desert base. An arbitrarily-given asset cannot be an essential constituent but only an accidental attribute of my person, for otherwise my identity would hang on a mere contingency, its continuity constantly vulnerable to transformation by

> experience, my status as a sovereign agent dependent on the conditions of my experience rather than epistemologically guaranteed. On Rawls' conception, no one can properly be said to possess anything, at least not in the strong, constitutive sense of possession necessary to the notion of desert.2

Sandel believes of Rawls that the claim about the moral arbitrariness of our attributes implies a metaphysical subject in which all empirical attributes are not only morally arbitrary but also metaphysically contingent. The metaphysical subject that is imputed to Rawls is in part the result of Sandel's refusal to accept that Rawls can successfully deny desert directly. Sandel argues that desert is not undermined by claiming that I cannot deserve what is arbitrarily given. He agrees with Feinberg that desert must presuppose some basis of desert that is not itself deserved.3 That basis is the possession of attributes. Sandel distinguishes possession in the strong sense from possession in the weak sense. The former seems to mean that one possesses one's attributes as essential properties to personal identity, whereas the latter appears to mean that one's empirical properties are contingent to personal identity.4 In other words, the possession of

---

2Sandel, pp. 85-86.
3Sandel, p. 84.
4In Sandel's discussion of Rawls' possible response to a meritocratic objection to the difference principle these meanings are revealed. Sandel interprets the meritocratic objection to the difference principle as independent of the question whether I deserve my attributes and as dependent upon whether my attributes are essential to personal identity. On page 74, Sandel states the meritocratic position as the following: "...genetic endowments are inviolable in a way that social or cultural characteristics are not, that a person's natural endowments are somehow more

characteristics is sufficient to provide a footing for desert. Hence, he contends that Rawls must go one step further and deny desert indirectly by arguing that the underlying subject does not have the requisite characteristics that could serve as a basis of desert. Sandel maintains that by appealing to a Kantian ego and claiming that all empirical properties are contingent Rawls can deny desert. If all empirical properties are contingent to personal identity I possess them in the weak sense. If I do not possess my attributes in the strong sense then there would be no foundation for desert because I cannot deserve what is not essential to my identity.5 A reconstruction of Sandel's argument will prove useful for exposing the fallacy. Sandel attributes the following view to Rawls:

I P1. Natural facts are neither deserved nor undeserved.
P2. Whatever is neither deserved nor undeserved is morally arbitrary.
P3. Our assets are natural facts.

---

essentially his, more deeply constitutive of his identity than his socially-conditioned attributes." More importantly, Sandel suggests that Rawls can respond to this objection by claiming that all empirical attributes are contingent. On the same page Sandel states, "Rawls objection would be stronger still, for his theory of the person implies that no characteristic, whether social or natural, can be essential in this way." Given Sandel's interpretation of the Rawlsian self he believes that the meritocratic must show that "...I possess my intelligence (in some non-arbitrary sense of possession),...".(page 77) Since Sandel's discussion concerns the relationship of our attributes to personal identity, it is plausible to interpret possession in the strong sense as meaning that those attributes are essential to our identity. Alternatively, possession in the weak sense means that our attributes are contingent to personal identity. The important point here is that Sandel connects the notion of possession to the identity relation.

5Sandel, p. 84.

CI. Therefore our assets are morally arbitrary and neither deserved nor undeserved.

II P1. Whatever is the product of something which is neither deserved nor undeserved is itself neither deserved nor undeserved.
P2. Our assets are morally arbitrary and neither deserved nor undeserved.
CII. Benefits that are the result of our assets or attributes are not deserved.

Sandel then argues that,

III P1. Attributes and their products are not deserved because they are arbitrary.
P2. The Kantian ego is one in which all empirical attributes are contingent.
P3. Whatever is morally arbitrary is metaphysically contingent to personal identity.
CIII. Therefore all assets and their products are contingent.

IV P1. What is contingent is inessential to personal identity.
P2. All assets and their products are contingent.
CIV. Therefore, all assets and benefits are contingent to personal identity.

V P1. The properties that provide the basis for desert claims are essential properties to personal identity but contingent properties are not.
P2. All assets and benefits are contingent to personal identity.
CV. Therefore, assets and attributes cannot provide the basis for desert claims because they are contingent.

In this reconstruction the silent premise is P3 in argument III and the silent conclusion is CIII, namely that all assets and their products are contingent. First, in order for Sandel's final conclusion (i.e., that our assets and attributes cannot provide the basis of desert) to follow from premises P1 and P2 in argument V, our assets would have to be not only morally arbitrary but also metaphysically contingent (i.e., we would have to assume premise P3 in argument III and its conclusion at CIII). For instance, if

these assets are not also metaphysically contingent how could Sandel claim that these assets could not provide a desert base if possession in the strong sense is the requirement for desert? The fallacious nature of Sandel's argument is revealed through premise P1 of argument I. The arbitrariness of our assets merely represents Rawls' view that our assets and social position should play no role in questions of how liberties and social goods should be distributed. By appealing to the Kantian ego Sandel falsely infers from this that whatever is arbitrary is contingent. For Sandel's argument to follow thus far, either he must believe that "arbitrary" simply means metaphysical contingency or he must believe that if our assets are morally arbitrary they then must be metaphysically contingent. In the former case, Sandel incorrectly applies the term 'arbitrary' to mean metaphysical contingency. In the latter case, Sandel fallaciously derives a metaphysical conclusion from an ethical premise. It simply does not follow that if our assets and talents should play no role when deciding on the principles of justice that those assets must also be metaphysically contingent. However, this last claim requires further argument.

Rawls' proposition that our talents are morally arbitrary, is independent of any metaphysical claim about the criteria for personal identity. Certain properties may be either metaphysically essential or inessential and they may also to be morally arbitrary. The confusion might be

symptomatic of a misunderstanding of the phrase "morally arbitrary". If something is morally arbitrary there is no reason that a particular characteristic, under cases of justice, should determine a pattern of distribution for certain goods. The moral arbitrariness of talents is independent of their subsequent metaphysical status in regard to personal identity. It is plausible that Rawls does view some of our natural assets and all of our social assets as metaphysically contingent. However, when Rawls discusses the notion of arbitrariness he is concerned with moral arbitrariness not metaphysical contingency. For Rawls, the metaphysical status of our attributes is irrelevant and independent from the question whether they should play a role in reasoning about justice. If we assume that Rawls does view certain features as metaphysically contingent (though he may not) Sandel claims that it follows that no empirical criteria can serve as necessary conditions for personal identity. But this conclusion is unfounded unless you also assume the following false premise: if some of my empirical or natural features are metaphysically contingent then no empirical properties are essential to my personal identity. This argument has the same logical structure as the following argument.

(1) I have 5 dogs.

(2) 2 of those dogs are black.

(3)Therefore, all 5 dogs are black.

When the logical structure is revealed through this simple but fallacious argument it is clear that this metaphysical conception of the self does not follow from A Theory of Justice. Given that Rawls' claims about the arbitrariness of our assets are normative as opposed to metaphysical, there is no reason to suppose that all empirical properties are inessential for personal identity. If there is a metaphysical conception of the self underlying A Theory of Justice, appeals to his ethical claims will be of little use in revealing it. For these claims lack the metaphysical status to direct us to a preferred conception of the self. On the one hand, if a given asset is morally arbitrary then this is compatible with the same attribute being metaphysically essential to one's identity. But the ethical question as to whether certain goods should be distributed on the basis of characteristics remains unresolved. On the other hand, if Sandel believes that certain attributes are metaphysically contingent, then it may be true that they are not essential to your personal identity but it does not follow that (a)those assets are not also morally arbitrary or (b) that they are significant from a moral point of view.

For Sandel's view to follow, considering the subject from a normative point of view entails a particular characterization of the subject from a metaphysical point of view. This is the general mistake that Sandel makes. When we are discussing a subject we are concerned with a person. A person also has different capacities. However, we do and

ought to separate those capacities. For example, when we are trying to determine the necessary criteria for personal identity we are discussing metaphysics and the nature of personhood. Yet, when we are concerned with the questions of responsibility we are focusing on the person in his or her moral capacity. So even if we are talking about the same entity (i.e., a person) we are discussing distinct features of a person, and hence those properties are not identical. The independence of moral theory is represented by requiring a criterion of identity that can be explained by it. Hence, one expects a metaphysical view of the subject to support one's ethical view. But ethical theorists can choose the appropriate criterion of identity that support their theory. This view is clearly supported by Rawls himself when he argues that although

> ...philosophy of mind may establish conditions that any correct criterion must satisfy, none of the traditional doctrines are affected by these constraints, at least not so long as these doctrines are applied under the normal conditions of human life....(A) criterion of identity is tailored to the requirements of a particular moral view. To this extent, the variations among the criteria are not antecedent to moral theory but explained by it.6

Section 2. The Constitutive Conception of the Self: Metaphysical or Moral

In addition to the disembodied subject, Sandel attributes a second conception of the self to Rawls. This is labeled the constitutive conception or the

---

6John Rawls, "The Independence of Moral Theory", Presidential Address to the American Philosophical Association, Eastern Division, in Proceedings and Addresses of the American Philosophical Association, 48, p. 15.

intersubjective conception of the self and it is derived from the Rawlsian difference principle. The defining feature of this view is that the essential properties of the self are identified with the aspirations, values, and relationships that are endorsed by the community.7 The Kantian ego and the constitutive conception of the self are incompatible in two ways. First, the Kantian conception of the self is one in which all empirical properties are contingent. Second, this notion of the self applies to empirically individuated human beings, namely, persons. In contrast, the constitutive conception of the self is identified by empirical properties, but they are the properties of the community such as the community's values and goals. The constitutive conception of the self "...may embrace more than a single empirically-individuated human being."8 The appropriate notion of the self is the community. Hence, Sandel not only believes that Rawls is committed to a metaphysical conception of the transcendental self but that he is also committed to the constitutive subject which is incompatible with the first view.
This conception of the subject involves a view that a member of the community is so attached to the community's values and goals that his identity is affected. Public life is an essential constituent of our identity, that is,"...the identity as well as the interests of the participants could

---

7Sandel, p. 62.
8Sandel, p. 80.

be at stake."9 Ultimately, in such an association my identity is defined by my attachments and commitments to the community.10 To say that a sense of the community is constitutive of me is to say that I identify my essential properties as those that are reflected in the community.

It is difficult to determine what Sandel's own view of the self is but there are clues given throughout the book. First, his aim is to give a "...philosophical anthropology in the broadest sense; philosophical in that it is arrived at reflectively rather than by empirical generalization, anthropology in that it concerns the nature of the human subject in its various forms of identity."11 It is puzzling to think of a person having various forms of identity if the issue under consideration is personal identity. Personal identity is concerned with identifying necessary and sufficient conditions for identifying the self as the same self over time. Hence, in our search for such a criterion we hope to be talking about the same thing to which this criterion will apply. The human subject or the person is the "thing" we should be considering. However, it is not at all clear that Sandel is talking about persons because he does not feel compelled to consider the possibility that certain physical characteristics, such as having a body and a brain are essential to our identity.

9Sandel, p. 62.
10Sandel, p. 179.
11Sandel, p. 50.

Sandel believes that the subject has various forms of identity which are revealed through reflection or self-understanding as opposed to empirical generalization. In addition to the intersubjective conception of the self Sandel introduces the intrasubjective conception of the self. This concept is similar to the intersubjective conception insofar as it is not restricted to an empirically individuated human being; and it seems to be an ethical concept. For instance, intrasubjective conceptions of the self

> ...allow that for certain purposes, the appropriate description of the moral subject may refer to a plurality of selves within a single, individual human being, as when we account for inner deliberation in terms of the pull of competing identities, or a moment of introspection in terms of occluded self-knowledge, or when we absolve some one from responsibility for the heretical beliefs 'he' held before his religious conversion.12

As in the case of the intersubjective conception of the self, the intrasubjective conception of the self may be of practical use if we are describing situations where this concept is viewed normatively. However, as a normative concept it does not involve metaphysical considerations. Here Sandel is not concerned with personal identity, but rather he is considering various ways a moral subject may be viewed. In short, from a metaphysical point of view of the subject the concern is with issues such as personal identity, whereas Sandel seems to be concerned with the "moral nature" of human beings that may reflect a more or

---

12Sandel, p. 62-63.

less expansive understanding of moral agency. Yet, this is not the case simply. We cannot ignore the fact that, when Sandel discusses the constitutive conception of the self, he describes it from a metaphysical stance. For instance, Sandel describes this view as one in which the community possesses attributes in the strong sense and as we have seen, possession here seems to mean possessing features as if they are essential to personal identity.13 In addition, even though the intersubjective conception of the self seems to be normative the constitutive conception of the self is often used interchangably with the intersubjective conception of the self.

There is some evidence that the constitutive and the intersubjective conceptions are intended to be one and the same. When discussing how Rawls can defend the difference principle, Sandel suggests that treating assets as common assets may be justified "...by allowing that, in certain moral circumstances, the relevant description of the self may embrace more than a single empirically-individuated human being....It appeals, in short to the intersubjective conception of the self."14 Sandel goes on to identify this conception of the self with the constitutive conception. He states that, "if the difference principle is to avoid using some as means to others' ends, it can only be possible under circumstances where the subject of possession is a 'we' rather than an 'I', which circumstances imply in turn the

13Refer to footnote 5 in this chapter on meritocracy.
14Sandel, p. 80.

existence of a community in the constitutive sense."15 From these statements it should be clear that Sandel often applies these terms as if they are the same concept. Yet, we have seen from the discussion of arguments I-V how Sandel thinks the view that assets and attributes cannot form the basis of desert is derived by Rawls indirectly from a metaphysical conception of the self. This suggests that the constitutive conception is also a metaphysical notion. This is clearly the case if possession is the undeserved basis for desert. Only by providing the appropriate metaphysical foundation does Sandel believe that Rawls avoids the difficulty of treating persons as means as opposed to ends in themselves. However, the intersubjective view of the self is referred to as appropriate in certain moral circumstances. Hence, that it is a normative notion.

I suggest that first we view the two conceptions of the self from a moral point of view. Considering them thus, we are not concerned with metaphysics because this issue is irrelevant to the issue at hand. By employing the intersubjective conception we may affirm an obligation to an artificial entity like a family or community as opposed to a particular human being. By employing the intrasubjective conception we may account for our inner deliberation between desires we can anticipate having in the future that may be incompatible with our present desires. This situation can be described in terms of a conflict between present and

---

15Ibid.

future selves.16 For Sandel, these conceptions are metaphors and are useful for describing the relevant agent in particular moral situations.17 Yet, if these concepts are to be viewed merely as useful metaphors, in an ordinary sense, there is no compelling reason to think that Rawls rejects them. Surely this kind of discourse is as useful to Rawls as it is to Sandel, as long as these terms are considered as strictly metaphorical. In fact, in an article Rawls allows for the term persons to

> on some occasions...mean human individuals, but in others it may refer to nations, provinces, businesses...and so on. The principles of justice apply to conflicting claims made by persons of all of these kinds. There is, perhaps, a certain logical priority to the case of human individuals: it may be possible to analyze the actions of so-called artificial persons as logical constructions of the actions of human persons, and it is plausible to maintain that the world of institutions is derived solely from the benefits they bring to human individuals.18

Considered normatively, the constitutive conception of the self is one in which for example, actions of a nation or company is the logical sum of the actions of its members. For instance, Rawls can claim that a company is morally responsible for a defective product that causes injury to an individual even though an individual designed that product. This description is plausible because it is based on the fact that in designing that product the individual is acting

---

16The conflict of desires that drug addicts may encounter might be a situation where a conflict of selves within a self is a useful application of this metaphor from a normative point of view.
17Sandel, p. 63.
18"Justice as Reciprocity", pp. 295-296.

as a representative of the company. And the moral responsibility lies with the company to oversee the acts of its representatives. The individual is prior however, because the act can always be traced to individual human beings. Rawls might apply this term in the sense described but he cannot allow the community to be prior to its members in that the community dictates one conception of the good to all its members. In short, for Rawls, a communitarian view cannot serve as a political ideal. Social facts are reducible to facts about individuals for Rawls.

Sandel himself suggests that even A Theory of Justice includes artificial entities when communities and social unions are discussed. For example, Rawls claims that "only in the social union is the individual complete", and "the self is realized in the activities of many selves".19 These statements represent the inter-personal view of the self. However, as a useful metaphor for describing an ethical agent it is compatible with the claims that some empirical properties are contingent or that all are contingent.

Similarily, the intrasubjective conception may be a useful way of describing our present and future desires or actions. Sandel correctly reveals that Rawls' appeals to the intrasubjective conception when, in his discussion of deliberative rationality, Rawls claims that "One who rejects equally the claims of his future self and the interests of others is not only irresponsible with respect to them but in

---

19Sandel, p. 81.

regard to his own person as well."20 But Sandel further claims that Rawls must stop short of accepting this conception because a person must be viewed as an enduring individual.21 However, if this concept is an ordinary metaphor and is not meant to replace our ordinary conception of a person it is unclear why the fact that a person must be viewed as an enduring individual affects the use of this concept in practical moral situations as a metaphor. I suspect that Rawls must claim that facts about our subselves as artificial entities constitute facts about ourselves as individuals which is analogous to the claim that social facts are reducible to facts about individuals.

Viewed as a useful normative metaphor the intersubjective conception does not provide the metaphysical clout to avoid either the objection of treating persons as means or the disembodied self. On Sandel's own interpretation of the problem, only a metaphysical conception of the self can serve as an opposing view to another metaphysical conception. If the intersubjective view of the self is metaphysical it does serve as an alternative. However, it is not viable because if it is metaphysical then my previous arguments for rejecting this conception stand as argued. Next I examine in more detail the metaphysical absurdity of replacing our ordinary conception of the self with the constitutive conception and the intrasubjective conceptions of the self.

20A Theory of Justice, p. 423.
21Sandel, p. 63.

Once the distinction between the subject from a metaphysical and normative point of view is pressed, the descriptions as formulated by Sandel lead to metaphysical absurdities. Sandel claims that,

> Intersubjective conceptions allow that in certain moral circumstances, the relevant description of the self may embrace more than a single, individual human being, as when we attribute responsibility or affirm an obligation to a family or community. ...Intrasubjective conceptions on the other hand, allow that for certain purposes, the appropriate description of the moral subject may refer to a plurality of selves within a single, individual human being, as when we account for inner deliberation in terms of the pull of competing identities, or moments of introspection in terms of occluded self-knowledge or when we absolve someone from responsibility for the heretical beliefs 'he' held before his religious conversion...to speak of selves within a ... self is not merely metaphorical but sometimes of genuine moral and practical import.22

For instance, consider the intersubjective self from a strict metaphysical point of view. I can die in several ways. I can die by selling my shares because that relation to my company is essential to my identity. Yet I am to find solace in the fact that being shot through the heart does me no harm because empirical attributes such as my physical body are not essential to my existence. As an individual person I am not properly viewed as a self but rather, the self is defined by the community to which I belong. The intrasubjective view of the self taken as a metaphysical view is equally absurd. On this view I can die several times because I must constantly make room for my

---

22Sandel, pp. 62-63.

future selves. Only by viewing these conceptions from a moral stance can we hope to avoid metaphysical absurdity.[23]

However, unless Sandel employs these concepts metaphysically then he has no objection against Rawls because there is no reason why Rawls would reject this way of talking. Conversely, if Sandel does apply these terms metaphysically they lead to absurd metaphysical conclusions. It seems that the normative view of community that Sandel advocates is transformed into a metaphysical view. Only from a normative point of view is Sandel's notion of community prima facie plausible. Taken metaphysically there is an additional problem with Sandel's argument because the constitutive conception of the self would be illegitimately derived from ethical claims. This is the topic of the next section.

Section 3. The Constitutive Conception Fallaciously Derived

The constitutive conception of the self is attributed to Rawls on the grounds that it naturally follows from his claim about treating our assets as if they are common assets. Furthermore, Sandel thinks that the Kantian self

---

[23]From a normative stance Rawls can allow for artificial entities such as communities, companies and the like. However, the important consideration here is not whether these artificial entities are viewed as having the same properties as its members but rather whether the company or the individual is either morally entitled to some benefit or morally responsible for some ill. This is be explained by whether an individual is acting as a representative of an artificial entity which allows us to explain this instance from a normative point of view independently of metaphysics. This also explains how social facts can be reduced to facts about individuals.

can serve as the vehicle whereby the constitutive conception of the self may be reached. For example, if all empirical properties are contingent to empirically individuated human beings then those properties cannot serve as necessary criteria for personal identity. If this is so, Rawls need only reject this view as an appropriate notion of the self. Instead, Rawls can, and according to Sandel should, expand the notion of the self to the community.

The constitutive conception is derived from two claims. The first is Sandel's claim that the Rawlsian view of arbitrariness commits him to a dispossessed subject. The second is Rawls' claim that we should treat our assets as common assets. The context of this derivation lies in Nozick's means-ends objection. Sandel agrees with Nozick and suggests that Rawls may affirm the difference principle as it stands but the difficulty is that he is open to the objection that the difference principle treats people as means as opposed to ends in themselves. The claim is that the difference principle treats people as means because it regards "...people's natural assets as common property...".24 Sandel suggests that Rawls can adopt a transcendental conception of the self and avoid the objection by claiming that "...not persons but only 'their' attributes are being used as means to others' well-being."25 But Sandel agrees with Nozick that this alternative is unacceptable because it leaves us "...with a subject so

24Sandel, p. 78.
25Ibid.

shorn of empirically-identifiable characteristics (so 'purified', in Nozick's word), as to resemble after all the Kantian transcendent or disembodied subject Rawls set out to avoid."26 The third alternative, unseen by Nozick, is Sandel's answer on behalf of Rawls, for avoiding dubious Kantian metaphysics and responding to the dilemma. It is a metaphysical conception of the self that denies that one is being used as a means rather than an end when one's assets are treated as common assets,

> ...not by claiming that my assets rather than person are being used, but instead by questioning the sense in which those who share in 'my' assets are properly described as 'others'. Where the first defense presses the distinction between the self and its attributes, the second qualifies the distinction between the self and the other by allowing that, in certain moral circumstances, the relevant description of the self may embrace more than a single empirically-individuated human being.27

Rawls' theory is open to this widening of the self because Sandel believes that Rawls endorses a theory of the person in which, "...all endowments are contingent and in principle detachable from the self, whose priority is assured...".28 Hence, Sandel advises Rawls to adopt this wider notion of the self to avoid lapsing into a transcendental self. This, according to Sandel, avoids both Kant's metaphysically suspect view and the charge of treating persons as means as opposed to ends.

The specific argument that Sandel employs to arrive at the constitutive conception of the self is as follows: The

---

26Sandel, p. 79.
27Sandel, pp. 79-80.
28Sandel, p. 79.

"...difference principle represents, in effect, an agreement to regard the distribution of natural talents as a common asset and to share in the benefits of this distribution whatever it turns out to be...".29 For Rawls, these talents are natural facts that are neither deserved nor undeserved and hence, are arbitrary from a moral point of view. However, Sandel then claims,

> in this way the difference principle acknowledges the arbitrariness of fortune by asserting that I am not really the owner but merely the guardian or repository of the talents and capacities that happen to reside in me, and as such have no special moral claim on the fruits of their exercise.30

Sandel contends that only by adopting the wider notion of the subject "...where the subject of possession is a 'we' rather than an 'I' ..." can Rawls avoid the Kantian transcendental self and the means-ends objection. In arguments I-V I show how Sandel arrives at the claim that our assets are contingent and hence cannot be the basis of desert but a formal reconstruction of Sandel's argument for the constitutive conception might prove useful.

1. All assets are contingent to the Kantian self.

2. If they are contingent they are inessential to personal identity.

3. If they are accidental to my identity I do not own or possess those assets in a strong sense.

4. Even though assets are inessential to the Kantian self qua individual it is possible that they are both owned and essential properties of some other self.

5. By adopting the constitutive self those

29Sandel, p. 70.
30Ibid.

properties may be regarded as essential and owned by the self qua community.

6. The basis necessary to make claims of desert is possession in the strong sense (i.e., those attributes must be viewed as essential to personal identity).

Therefore the community deserves individual assets and their benefits because they are possessed as essential properties by the community.

I argue in section 1 of this chapter that Sandel believes that desert has a basis that is itself not deserved. Hence, it follows that Sandel also accepts that if we, as subjects, possessed our attributes as essential to our personal identity then this would be sufficient for also claiming that we deserve the benefits that flow from those assets. Moreover, we deserve those benefits because metaphysically those assets are essential constituents of our identities. Sandel believes that Rawls must deny desert indirectly and accomplishes this by appealing to the Kantian subject. However, by employing a similar tactic in the argument for the constitutive conception of the self Sandel believes that Rawls can avoid dubious Kantian metaphysics and still be equipped to deny desert indirectly. The wider notion of the self provides the metaphysical foundation for denying desert by restricting the application of a concept of the self to the community. The Kantian self is overridden because in this case all empirical properties are contingent and hence cannot be owned by the self qua individual. Hence, desert is denied qua individual by claiming that if the community is viewed as the proper notion of the self then the

community metaphysically possesses the assets in a sense necessary for desert. Altering our concept of the self means that the individual as such does not even enter the discussion of desert because desert is only a concern for selves. And the community deserves those assets and benefits because the individual is denied any such status. If the constitutive conception of the self is to serve as a replacement for the Kantian self it seems that it must be considered from a metaphysical point of view. The metaphysical status of the constitutive conception is grounded in the fact that the basis of desert depends on ownership, which is in turn determined via essential properties. If it is metaphysical then as a conclusion it is fallaciously derived from ethical premises, as is the transcendental conception of the self. For Rawls, arbitrariness is normative and does not imply any particular metaphysical subject.

Section 4. Rawls' Reply to the Dilemma

If neither conception of the self taken metaphysically is a reasonable alternative Rawls must show how his account avoids treating individuals as means. I believe that this may be answered on normative grounds without relying on metaphysical considerations. First, Rawls makes it clear that the reason why the difference principle does not treat people as means is based on the fact that it is one of the principles that was chosen under the conditions that treat individuals as free and equal. Hence, "...treating men as

ends in themselves implies at the very least treating them in accordance with the principles to which they would consent in the original position of equality."[31] In other words, to treat one as a "self-originating source of valid claims"[32] is to treat one as an end in himself.

A second reply is found in the desert argument. For Rawls a person's talents and capacities are morally arbitrary because they are the result of the natural lottery and hence no one is in a position to make claims of desert for the benefits that result from those attributes. Those attributes are natural facts, and hence any pre-institutional notion of desert is inapplicable. Yet, Rawls believes that what is just or unjust is how institutions deal with these facts. The benefits that result from our attributes cannot be said to be deserved hence those attributes in turn, cannot provide any claims to having a right to a particular distribution of benefits. If this is so, it is unclear why treating these assets as common assets should be viewed as treating people as means to others' ends. I am suggesting that this is true because one must have a right that society value a particular virtue which determines a particular distribution of goods before the difference principle treats people as means as opposed to ends. It seems that for Rawls such a right cannot be established because desert cannot be applied antecedent to

---

31 A Theory of Justice, p. 180.
32 John Rawls, "Kantian Constructivism in Moral Theory," The 1980 John Dewey Lectures at Columbia University in The Journal of Philosophy, vol. 77, p. 543.

institutions. In short, if these rights cannot be established without presupposing desert prior to institutions, then individuals cannot complain that they are being treated unfairly, because where there are no rights there can be no violation of rights. The important point for Rawls is that the morally significant consideration in treating people as ends is that we respect the choices they make. Moreover, on this account the metaphysical status of persons is irrelevant to the means-ends distinction.

It must be noted that this argument only affirms that no one deserves his or her particular talents. But this does not imply alternatively, that the reason why assets are treated as common assets is because society deserves those assets. However, Sandel believes that unless this proposition can be established "there would seem no grounds for favoring a utilitarian dispensation of such assets and endowments rather than just letting them lie where they fall."33 By examining the justification for the difference principle in terms of the fairness-ensuring description of the original position I show that Rawls need not argue that society either owns or has some claim to those assets in order to argue that they cannot let those assets lie where they fall.

If the original position is an instance of pure procedural justice that is fair, then the agreements reached in it are also fair. If the morally arbitrary character of

---

33Sandel, p.141.

our assets is accepted as a premise then the description of the initial situation is also accepted because it is in part representative of that premise. Let us suppose that the parties to the original position accept the claim that one's assets and talents are arbitrary from a moral point of view. Given the constraints that are embedded in the reasoning process, such as the veil of ignorance and the maximin principle, it is plausible that the parties agree to treat these assets as common assets.34 But the interesting question is whether the premise that society has some claim to those assets plays any role in the argument. This may be answered in part by the view that the parties agree to treat assets as common assets, not because society has any claim to those assets, but rather because letting them lie where they fall is tantamount to denying the premise that our assets are in fact morally arbitrary. Letting these assets lie where they fall is in effect allowing the better endowed to be advantaged when the distribution of social goods is considered. The better endowed will have the necessary power to accumulate more of those goods than the less endowed. This in turn, is to deny the claim that our assets are morally arbitrary. Yet, this argument is independent of making good the claim that society has a right to those assets in a way in which the individual does not.

34The maximin principle restricts our reasoning for the two principles insofar as we are to choose the principles as if our worst enemy is to assign us our place in society.

Additional support for this view is found in a later article by Rawls called "Justice as Fairness: Political not Metaphysical". Rawls justifies the veil of ignorance which represents the claim that our assets are morally arbitrary. In effect it is also a justification for why we cannot let those attributes lie where they fall. Rawls claims that the

> ...reason why the original position must abstract from and not be affected by the contingencies of the social world is that the conditions for a fair agreement on the principles of political justice between free and equal persons must eliminate the bargaining advantages which inevitably arise within background institutions of any society as the result of cumulative social, historical, and natural tendencies.35

Thus, if we let assets lie where they fall then the better endowed will inevitably occupy an advantageous bargaining position. This represents the view that our assets and endowments are morally relevant in the distribution of social goods. Yet this is the very claim that Rawls denies. In the next section I wish to discuss a fundamental difference between Sandel and Rawls. Sandel's vision of communitarianism must presuppose an objective moral realm. Rawls, on the other hand, is not committed to one. In terms of trying to solve the practical problem of social justice I argue that Rawls' theory is the more plausible insofar as the assumptions that he makes are much weaker and hence, make the resolution of this problem feasible.

35"Justice as Fairness: Political not Metaphysical", pp. 235-236.

Section 5. Objectivism-The Difference between Rawls and Sandel

Sandel endorses the view that certain values such as benevolence and altruism may be so strongly held that, if they are denied expression at a public level as first virtues, this is tantamount to denying one's identity. In other words, Sandel believes that the possibility exists that one may not be able to pursue one's commitment to benevolence and altruism fully if it is not pursued as a political ideal. Although this view is not explicitly stated by Sandel it is at least indirectly implied for he objects to Rawls on the grounds that this is denied as a possibility. We must remind ourselves that these values are not chosen for Sandel but rather they are discovered to comprise our essence. If they are discovered they are in some sense "given". We are required to identify certain values or aims as essential and upon this discovery we have revealed the values that ought to be upheld. Yet, this requires that there is an objective moral realm that merely awaits discovery. Furthermore, regardless whether individuals either accept or correctly identify which values and beliefs define their identities, they nonetheless exist independently of individuals beliefs. In other words, there is a true or correct moral realm that exists independently of our capacity to either accept or identify it. This is a realm in which certain values are standards that are representative of values that are morally good independently of whether anyone believes this to be true or not. However,

if communitarian values such as benevolence and altruism are to provide the standards for moral behavior and also pursued as a political ideal then all incompatible conceptions of the good are excluded.

Rawls' account is more plausible because he is not committed to an objectivist or subjectivist account of morality. Rawls argues that unjust conceptions of the good (i.e., those that conflict with the two principles of justice) must be rejected. This is so, not because A Theory of Justice is the true moral theory which represents an objective moral realm but rather because conceptions of the good should play no role in determining how liberties and social goods should be distributed under cases of justice. In other words, conceptions of the good cannot be the basis for a political conception of justice. In terms of the good, there is no agreement on which doctrine should be advanced and hence, for Rawls, advancing one cannot be accomplished without authoritarian means. Since doing so is generally undesirable we must allow for the most comprehensive set of conceptions of the good. Given that Rawls is concerned with solving the problem of social justice in a modern constitutional democracy the fact that individuals do not agree on conceptions of the good poses a difficult challenge. If we do not agree on these questions, how can principles be devised that all can accept? Rawls suggests that

> ...whenever a sufficient basis for agreement among citizens is not presently known. or recognized, the

> task of justifying a conception of justice becomes: how can people settle on a conception of justice, to serve this social role, that is (most) reasonable for them in virtue of how they conceive of their persons and construe the general features of social cooperation among people so regarded?36

For Rawls, a conception of justice is in some sense objective but only insofar as the principles of justice are grounded in agreement that can be objectively determined. In contrast to Sandel's position,

> what justifies a conception of justice is not its being true to an order antecedent to and given to us, but its congruence with our deeper understanding of ourselves and our aspirations, and our realization that, given our history and the traditions embedded in our public life, it is the most reasonable doctrine for us....Kantian constructivism holds that moral objectivity is to be understood in terms of a suitably constructed social point of view that all can accept.37

Therefore, the only moral facts that are accepted are those that are publicly recognized by rational agents that are represented as free and equal.

Rawls' project is defined by certain conditions. First, if our concern is social justice in a democracy our conception "...must allow for the diversity of doctrines and plurality of conflicting and indeed incommensurable, conceptions of the good ...".38 Hence, conceptions of the good cannot provide the basis for the two principles. Rawls' project is to articulate what we do agree upon vis-a-vis the two principles of justice.

36"Kantian Constructivism in Moral Theory", p. 517.
37Ibid. p. 519.
38"Justice as Fairness: Political not Metaphysical", p. 225.

Rawls believes that the differences in conceptions of the good and disputed moral, religious and philosophical questions are too important to be abandoned and he realizes they cannot be resolved at a public level without introducing some view of enforced re-education or the autocratic use of power.39 He contends that, "...philosophy as the search for truth about an independent metaphysical and moral order cannot,..., provide a workable and shared basis for a political conception of justice in a democratic society."40 As a practical solution to social justice Rawls deliberately avoids these considerations because the autocratic use of power is contrary to our considered convictions about justice. Since we do not view this as a viable option Rawls' solution seems more plausible than Sandel's because we do not agree on what conception of the good should be affirmed. In the end, people must agree to disagree on substantive conceptions of the good. Yet, they can agree to the most reasonable understanding of justice such that the principles of justice express respect for their most deeply felt and perhaps incommensurable philosophical, religious and moral views.

I believe that Rawls rejects perfectionism in much the same way he would reject Sandel's position. He considers two variants of perfectionism. The first is concerned with "...directing society to arrange institutions and to define the duties and obligations of individuals so as to maximize

---

39Ibid, p. 231.
40Ibid, p. 230.

the achievement of human excellence in art, science, and culture."[41] The second variant asserts that the principle of perfectionism is accepted as one out of several standards in an intuitionists theory.[42] Hence, here perfectionism is not pursued absolutely but rather is balanced against other principles. Nonetheless, both variants are rejected on the grounds that the principle of perfection implies some objective standard of human excellence and if it is adopted as a political principle it requires all citizens to accept one conception of the good. But the people in the original position know that they may have certain moral and religious interests and cultural ends that they cannot jeopardize.[43] This represents the fact that in modern constitutional democracies individuals do not have

> ...an agreed criterion of perfection that can be used as a principle for choosing between institutions. To acknowledge any such standard would be, in effect, to accept a principle that might lead to a lesser religious or other liberty, if not the loss of freedom altogether to advance many of one's spiritual ends.[44]

Even though perfectionism is denied as a political ideal individuals can evaluate the intrinsic worth of their desires. However, the pursuit of these human excellences is limited by the principle of free association.[45] The important consideration is that although liberties and social goods cannot be distributed according to virtue,

---

41A Theory of Justice, p. 325.
42Ibid.
43Ibid, p. 327.
44Ibid.
45Ibid, p. 328.

moral worth or human excellence this does not entail that individuals are incapable of evaluating their lives or their desires on the basis of intrinsic worth.

Communitarianism is similar to perfectionism in that it promotes particular virtues as human excellences. Values such as altruism and benevolence are the mainstay of this position. But even though one may be committed to benevolence and altruism to the extent that one feels that it should be our political ideal this cannot provide a practical solution to social justice in a pluralistic democracy. If one is committed to one's values one must accept a principle of tolerance because there may be others that are equally committed to different conceptions of the good. Therefore, the pursuit of conceptions of the good must be limited in insofar as, liberties and freedom to pursue one's conception of the good are given to all. Despite this limitation the two principles of justice provide the conditions that encourage the development of associations and communities that represent different and incommensurable conceptions of the good.

At this point it might be argued that Rawls' position is merely a modus vivendi. In other words, the objection is that the agreement reached by the people in the original position is merely a working arrangement based on self- or group-interests. We agree to the two principles but it is claimed that our agreement is a compromise.46 Rawls has

---

46"The Idea of an Overlapping Consensus", p. 10.

several answers to this objection. First, Rawls claims that if the modus vivendi objection is motivated in the hopes that a conception of justice can be found that embodies a general and comprehensive conception of the good, then this is impossible. This of course depends on Rawls' claims that pluralism is a fact and that the authoritarian use of power is rejected.47

The second reason that this is not merely a modus vivendi position is central to the fact that the object of consensus namely, our conception of justice, is a moral conception that is based on moral grounds.48 "An overlapping consensus, therefore, is not merely a consensus on accepting certain authorities, or on complying with certain institutional arrangements, founded on a convergence of self- or group-interests."49

The third reason why this is not a mere agreement is represented by Rawls by the feature of stability. In virtue of the fact that all those with opposing moral, religious and philosophical views agree to this conception of justice on the basis that this represents an overlapping consensus of their views, Rawls' conception of justice is stable unlike a modus vivendi. Because "...each view supports the conception for its own sake, or on its own merits;..."50 its stability does not seem to depend on the distribution of power amongst those with different conceptions of the good.

47Ibid.
48Ibid, p. 11.
49Ibid.
50Ibid.

However, in a case which seems to represent a modus vivendi view this is not true. Rawls uses the example of the views between Catholics and Protestants in the sixteenth century. In this case he claims that there is no consensus on a principle of toleration because the citizens believe that it is the duty of those in power to promote the true faith. If neither faith is dominant then the agreement to a principle of toleration is a modus vivendi because if the situation changes and either faith becomes dominant the principle of toleration would be rejected. "Stability with respect to the distribution of power no longer holds".51 From this it is clear that Rawls' position provides a practical solution to the problem of social justice without lapsing in to a position that is merely a modus vivendi.

The last chapter provides reasons for thinking that Rawls is not committed to either conception of the self. Moreover, it shows the way in which Sandel confuses issues of advocacy and issues of metaphysics. And by so doing Sandel jeopardizes the prima facie plausibility of his own vision of society. I leave many aspects of Sandel's theory untouched but I show how Sandel's attribution of the conceptions of the self are not entailed by Rawls' normative theory.

Chapter I was designed as the first stage of a three part argument in the thesis. A cloud of suspicion is cast on the view of the Kantian self that Sandel attributes to

---

51Ibid.

Rawls once Sandel's systematic misunderstanding of A Theory of Justice is revealed. Sandel's reading of justice as the first virtue of institutions, I argued, is mistaken. Sandel claims that for Rawls the application of the principles of justice and conflict are necessary conditions for justice to remain the first virtue of institutions. However, Rawls need only claim that these conditions are sufficient but not necessary. I have shown two of three arguments Sandel presents in defense of his claim that Rawls must presuppose a Kantian ego are also based on a misreading of Rawls. Sandel is convinced that if the principles of justice are derived independently from conceptions of the good, and the justification for procedure must be prior to the justification of the principles themselves then the Kantian self is presupposed. Sandel errs, I have argued, in assuming that if the principles of justice cannot be based on conceptions of the good life then no societal values can be assumed. I have established Rawls' commitment to the claim that not all values can be assumed but I have further shown that he is not committed to the claim that no values can be assumed. Sandel's second mistake is assuming that the justification for Rawls' procedure must come wholly prior to the principles themselves. Rawls' coherence theory of justification undercuts both objections. In the first case, some views of justice that upon reflection are accepted will be assumed. In the second case, the justification for the procedure is not wholly prior to the

principles because it depends, in part, on whether the principles of justice cohere with our considered judgements of justice. The justification of procedure is connected to the content of the principles by appealing to how well the principles match our considered judgements about justice.

The second part of my argument was presented in chapter II. The doubts raised in chapter I were insufficient to discard Sandel's criticism of Rawls outright. It could have been the case that even on Sandel's misreading of Rawls, the Kantian ego must be presupposed by A Theory of Justice. In chapter II this possibility is eliminated. Here, I suggested three different interpretations of the Kantian self that could have vindicated Sandel's critique of Rawls. Although the three interpretations of the transcendental self do indeed logically follow from Rawls' theory I established that none of these views are options for Sandel because they lack the metaphysically problematic character that provides his central objection against Rawls. Hence none of the conceptions of the self that follow from A Theory of Justice can support Sandel's criticism of Rawls. I also have suggested a fourth interpretation of the transcendental self that fulfills the precondition of being metaphysically problematic which I argued is adopted by Sandel. This conception of the self is one in which all empirical properties are contingent to personal identity. Moreover, I have shown that Sandel's interpretation is based on a misreading of Kant. In virtue of the fact that Sandel

equates the Rawlsian self with the Kantian self Sandel's interpretation of Rawlsian metaphysics was also revealed. Finally, I have exposed the fallacious character of Sandel's arguments by showing how he both conflates issues concerning the epistemology of the self with issues concerning the metaphysics of the self and issues concerning personal identity with issues concerning individuation.

In chapter III the third and final part of the argument was presented. Here, I argued that the constitutive conception of the self must be taken from a metaphysical point of view. With Sandel's preferred interpretations of both the Kantian ego and the constitutive conception of the self in hand I have shown why they do not follow from Rawls' view. Both views of the self represent metaphysical propositions that are illegitimately derived from ethical propositions hence they are based on unsound arguments. Central to this argument is Sandel's claim that Rawls can avoid the charge of treating persons as means as opposed to ends, and retain his view of distributive justice by either adopting the Kantian self or the constitutive conception of the self. Sandel believes that the first alternative is unacceptable because it is metaphysically dubious. The latter alternative is more palatable to Sandel's communitarian tastes but, as a liberal, Rawls would reject this because it entails abandoning his individualistic project. I provided a reading of A Theory of Justice, unseen by Sandel, which does not depend on metaphysical

conceptions of the self. Rawls avoids the charge of treating individuals as means in virtue of the fact that to treat one as an end in himself is to treat him as the author of his own laws. Finally, I demonstrated that Sandel's view of society is less plausible than Rawls' view insofar as Sandel must assume the existence of an objective moral realm. Rawls on the other hand, need not assume an objectivist or a subjectivist account of morality.

BIBLIOGRAPHY

Allison, Henry. Kant's Transcendental Idealism: An Interpretation and Defense. London: Yale University Press, 1983.

Ameriks, Karl. Kant's Theory of Mind. Oxford: Claredon Press, 1982.

Beck, Lewis. A Commentary on Kant's Critique of Practical Reason. Chicago: The University of Chicago Press, 1960.

Cooper, Wes. "The Perfectly Just Society." In Philosophy and Phenomenological Research. Vol. 38, 1978, 46-55.

Doppelt, Gerald. "Rawls' Kantian Ideal and the Viability of Modern Liberalism." In Inquiry. Ed., Alastair Hannay. Norwegian University Press. Vol. 31, No. 4, December 1988.

Hampton, Jean. "Contracts and Choices: Does Rawls Have a Social Contract Theory?" In The Journal of Philosophy. Vol. LXXVII, NO. 6, June 1980, 315-338.

Kant, Immanuel. "Grounding for the Metaphysics of Morals." In Ethical Philosophy. Translated by James Ellington. Cambridge: Hackett Publishing Company, 1983.

Kant, Immanuel. Critique of Judgement. Translated by J.H. Bernard. New York: Hafner Publishing Company, 1972.

Kant, Immanuel. Critique of Pure Reason. Ed., Norman Kemp Smith. London: Macmillan, 1968.

Kant, Immanuel. "Perpetual Peace: A Philosophical Sketch." In Kant's Political Writings. Ed., Hans Reiss. Cambridge: Cambridge University Press, 1970, 93-103.

Kitcher, Patricia. "Kant's Paralogisms." In The Philosophical Review. Vol. XCI, No. 4, 1982, 515-547.

Kripke, Saul. Naming and Necessity. Oxford: Basil Blackwell, 1980.

Kymlicka, William. "Liberalism and Communitarianism." In Canadian Journal of Philosophy. Vol. 18, June 1988, 181-203.

Larmore, Charles. Patterns of Moral Complexity. Cambridge: Cambridge University Press, 1987.

Narveson, Jan. "Equality vs. Liberty: Advantage, Liberty." In Social Philosophy and Policy. 2:1, Autumn 1984, 33-60.

Nozick, Robert. Anarchy, State, and Utopia. New York: Basic Books, Inc., Publishers, 1974.

Parfit, Derek. Reasons and Persons. New York: Oxford University Press, 1986.

Rawls, John. A Theory of Justice. Cambridge: The Belknap Press of Harvard University Press, 1971.

Rawls, John. "Two Concepts of Rules." In The Philosophical Review. 64, 1955, 3-32.

Rawls, John. "The Sense Of Justice." In The Philosophical Review. 72, 1963, 281-305.

Rawls, John. "Distributive Justice." In Philosophy, Politics and Society. Third series, ed., by Peter Laslett and W.G. Ruciman: 58-82 London: Basil Blackwell, Ltd., 1967.

Rawls, John. "Justice as Reciprocity." In Choice and Action: An Introduction to Ethics, ed., by Charles Reid: 292-323. New York: Mcmillan Publishing Company, 1981.

Rawls, John. "The Independence of Moral Theory." Presidential Address to the American Philosophical Association, Eastern Division, 1974. Proceedings and Addresses of the American Philosophical Association 48, 1874-1975, 5-22.

Rawls, John. "A Kantian Conception of Equality." In Property, Profits, and Economic Justice, ed., by Virginia Held: 198-208. Belmont, California: Wadsworth Publishing Company, 1980.

Rawls, John. "Kantian Constructivism in Moral Theory." The 1980 John Dewey Lectures at Columbia University. In The Journal of Philosophy. Vol. 77, No. 9, 1980, 515-572.

Rawls, John. "Justice as Fairness: Political not Metaphysical." In Philosophy and Public Affairs. Vol. 14, No. 3. Summer 1985, 223-251.

Rawls, John. "The Idea of an Overlapping Consensus." In Oxford Journal of Legal Studies. Vol. 7, No. I, 1987, 1-25.

Sandel, Michael. Liberalism and the Limits of Justice. New York: Cambridge University Press, 1982.

Stahl, Gary. "Locating the Noumenal Self." In Kant-Studien. Vol. 72, 1981, 31-40.

Taylor, Charles. "What is Human Agency." In Human Agency and Language (Philosophical Papers Vol. 1). Cambridge: Cambridge University Press, 1985, 15-45.

Taylor, Charles. "The Concept of a Person." In Human Agency and Language (Philosophical Papers Vol. 1). Cambridge: Cambridge University Press, 1985, 77-97.

Taylor, Charles. "Interpretations and the Sciences of Man." In Philosophy and The Human Sciences (Philosophical Papers Vol. 2). Cambridge: Cambridge University Press, 1985, 15-58.

Taylor, Charles. "Kant's Theory of Freedom." In Philosophy and The Human Sciences (Philosophical Papers Vol. 2). Cambridge: Cambridge University Press, 19851, 318-339.

Werkmeister, W.H. Kant the Architectonic and Development of His Philosophy. Illinois: Open Court, 1980.

Wolff, R.P. Kant A Collection of Critical Essays. New York: Anchor Books, 1967.